مركز البحوث

تحليل البيانات الاقتصادية

تأليف

جاري كوب

ترجمة

د. فهد بن خلف البادي

راجع الترجمة

د. عبدالله بن صالح الحميد

بسم الله الرحمن الرحيم

مركز البحوث

تحليل البيانات الاقتصاديـــة

تأليف

جـــاري كـــوب

ترجمـــة

د. فهد بن خلف البـــادي

راجع الترجمـــة

د. عبدالله بن صالح الحميد

١٤٣٠هـ - ٢٠٠٩م

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

كوب، جاري

تحليل البيانات الاقتصادية/ جاري كوب،

فهد بن خلف البادي، عبدالله بن صالح الحميد

الرياض، ١٤٣٠هـ

٤٥٦ص؛ ١٧X٢٤سم

ردمك: ٩٧٨-٩٩٦٠-١٤-١٧٩-٤

١- التحليل الاقتصادي أ. البادي، فهد بن خلف (مترجم)

ب. الحميد، عبدالله بن صالح (مراجع) ج. العنوان

ديوي ٣٣٩ ١٤٣٠/٣٩٠٠

رقم الإيداع: ١٤٣٠/٣٩٠٠

ردمك: ٩٧٨-٩٩٦٠-١٤-١٧٩-٤

هذه ترجمة كتاب:

ANALYSIS OF ECONOMIC DATA

SECOND EDITION

By

Gary Koop

University of Leicester

John Wiley & Sons, Ltd

2005

المحتويـــــات

مقدمـــة المترجـــم

تنبع أهمية هذا الكتاب من الدور المحوري للبيانات الاقتصادية في وضع خطط التنمية، ورسم السياسات الاقتصادية، واتخاذ القرارات الاستثمارية على مستوى المؤسسات والأفراد، ومتابعة الأداء الاقتصادي عامة. كما أن تحليل البيانات الاقتصادية يساعد على سبر مدلولاتها ومن ثم يمكن الحصول على مؤشرات تمكننا من توقع التغيرات المستقبلية في النشاط الاقتصادي.

يهدف مؤلف هذا الكتاب إلى تيسير مادة الاقتصاد القياسي للطلاب والباحثين، وذلك من خلال تقديم الأساليب والأدوات والنماذج القياسية بأسلوب سهل وبعيد عن استخدام الاشتقاقات الرياضية المعقدة والنظريات الإحصائية، متوسعاً في استخدام الأمثلة والتمارين ليتيح للطلاب والباحثين تطبيق تلك الأساليب والنماذج القياسية بيسر وسهولة. كذلك يساعد الكتاب الطلاب والباحثين والمحللين الذين يعدون تقارير وبحوثاً ودراسات في مجال الاقتصاد وإدارة الأعمال.

ويتطلب فهم هذا الكتاب الإلمام بأساسيات استخدام برامج الحاسب الآلي الجاهزة مثل برنامج إكسل، كما أن النماذج القياسية تتطلب المعرفة باستخدام برامج الحاسب الآلي القياسية.

يُعد هذا الكتاب من أهم الكتب التي تهتم بفهم وتحليل البيانات الاقتصادية عن طريق التعرف على أساليب وأدوات تحليل البيانات، وطرح أمثلة وبيانات حقيقية تساعد الباحثين على تحليل النتائج القياسية والمؤشرات التي يتم الحصول عليها باستخدام تلك الأساليب والأدوات والنماذج القياسية.

ورغم أهمية هذا المجال المعرفي فإن المكتبة العربية تعاني نقصاً فيه، تأليفاً وترجمةً، وقد شجعني ذلك علي ترجمته في ظل وجود عقبات كثيرة، فتوخيت العبارات المباشرة السهلة التي لا ترهق القارئ دون الإخلال بالأفكار والمفاهيم

التي يتضمنها الكتاب، وكذلك اجتهدت في ترجمة بعض المصطلحات المستخدمة في هذا الكتاب التي لا يوجد لها مقابل محدد في اللغة العربية، واستشرت بعض زملاء المهنة والتخصص في ترجمة هذه المصطلحات حتى تمت ترجمتها بشكل واضح يساعد القارئ على فهمها. أرجو أن أكون قد وفقت في ترجمته، والله من وراء القصد.

المترجـــم

د. فهد بن خلف البادي

مقدمة الطبعة الأولى

يهدف هذا الكتاب إلى تعليم مادة الاقتصاد القياسي للطلاب الذين لا تكون هذه المادة متطلباً أساسياً لهم. وهم الطلاب الذين يودون تطبيق أساليب الاقتصاد القياسي بمعرفة كافية في يسر وسهولة، وذلك في سياق حل المسائل التطبيقية الواقعية. والكتاب يستهدف بالدرجة الأولى طلاب الجامعات الذين يمكنهم الاستفادة منه في التحليل القياسي للبيانات أو مرجعاً تطبيقياً في الاقتصاد القياسي يمكنهم الرجوع إليه في أي وقت.

وقد كان هذا الكتاب نتاج مقرر دراسي قمت بتدريسه في جامعة أدنبره تحت عنوان "تحليل البيانات الاقتصادية". وقبل بدء تدريس هذا المقرر بالجامعة المذكورة، كان يُطلب من جميع الطلاب دراسة مادة في الإحصاء والاحتمالات في السنة الأولى أو الثانية من دراستهم الجامعية. كذلك كان يُطلب من الطلاب المتخصصين في الاقتصاد دراسة مقرر في الاقتصاد القياسي في السنة الثالثة أو الرابعة. إلا أن ذلك لا يُطلب من الطلاب غير المتخصصين (مثل: طلاب الاقتصاد والعلوم السياسية أو طلاب الاقتصاد وإدارة الأعمال) على اعتبار أنهم دخلوا في سنواتهم الجامعية الأخيرة وبعدها مباشرة سوف يدخلون إلى سوق العمل، وفي جعبتهم مقرر دراسي أساسي واحد في الاحتمالات والإحصاء. هؤلاء الطلاب يكونون في الغالب غير مهيئين جيداً لتحليل البيانات الاقتصادية الحقيقية بإدراك كاف. إن مادة الاقتصاد القياسي هي مادة مهارة أساسية لمشاريع وبحوث التخرج الجامعية، ولكلية الدراسات العليا، وكما هو الحال لمعظم الفرص الوظيفية المتاحة للاقتصاديين، لذا تنامى الشعور بأن هناك حاجة إلى مقرر دراسي جديد يقدم أساساً عملياً قوياً في أدوات تحليل البيانات الاقتصادية. وكان ثمة اتفاق عام في قسم الاقتصاد بجامعة أدنبره على إضافة الأساسيات التالية عند إعداد المقرر الدراسي الجديد:

1-يجب أن يشمل معظم الأدوات والنماذج المستخدمة في البحوث القياسية الحديثة (مثل: الارتباط، الانحدار، وتطورات أساليب السلاسل الزمنية).

2-يجب أن يكون بصورة مبسطة وغير رياضية، وإنما يعتمد على الفهم اللفظي والبياني.

3-يجب أن يحتوي على استخدام شامل لأمثلة حقيقية للبيانات ويضع الطلاب في تجربة واقعية واستخدام حقيقي للحاسب الآلي.

4-يجب أن يكون مختصراً. وعلى كل، فإن الطلاب ولاسيما الطلاب الذين يدرسون تخصصات جامعية مشتركة (مثل: الاقتصاد وإدارة الأعمال، أو الاقتصاد والعلوم السياسية) يجب أن يكونوا ملمين بكم واسع من المواضيع. ونادراً ما يكون لدى الطلاب وقت أو رغبة في دراسة الاقتصاد القياسي بشكل متخصص.

وهذا الكتاب يتبع هذه المبادئ الأساسية. ويهدف إلى تعليم الطلاب أدوات الاقتصاد القياسي المنظورة بشكل مبسط وسلس، وذلك باستخدام الحدس غير الرياضي البسيط والأمثلة العملية. والأفكار المهيمنة على الكتاب هي المفاهيم المتعلقة بالانحدار والارتباط. هذه المفاهيم البسيطة تعد سهلة نسبياً في الحث على استخدام الحدس اللفظي والبياني، وتتضمن العديد من النماذج والأساليب المعقدة (مثل: تكامل المشترك وجذر الوحدة) المستخدمة في البحث الاقتصادي في الوقت الراهن. فإذا استطاع الطالب أو الطالبة استيعاب مفاهيم الارتباط والانحدار استيعاباً جيداً فإنه يمكنه أو يمكنها فهم وتطبيق الأساليب المستخدمة في الاقتصاد القياسي المتقدم والإحصاء.

وقد أنجز الكتاب ليكون ذا صلة بالحاسب الآلي واستخداماته؛ وذلك لقناعتي بأن الخبرة العملية في استخدام الحاسب الآلي المزودة بمحاضرات منهجية، هي

أفضل طريقة ليتعلم الطلاب مهارات تحليل البيانات عملياً. وقد تم توضيح كل النقاط النظرية الواردة في الكتاب بأمثلة اقتصادية عملية؛ حتى يتمكن الطالب من التوسع في استخدام الحاسب الآلي. وأعتقد اعتقاداً قوياً أن كل ساعة يقضيها الطالب أمام الحاسب الآلي تعادل عدة ساعات يقضيها الطالب في محاضرة.

ويتميز الكتاب بأنه سهل التناول من قبل الطلاب؛ وذلك لأنه يشتمل على قدر قليل من الرياضيات. وفيما عدا الملاحق الإضافية في بعض أجزائه، فإن هذا الكتاب لا يفترض أي خلفية في مادة الرياضيات بعد مرحلة ما قبل الجامعة.

أما عن الطلاب غير الملمين بهذه الأساسيات (مثل: معادلة الخط المستقيم، عمليات الجمع، اللوغاريثمات) فإن هناك العديد من الكتب التي يمكن أن تقدم خلفية كافية عنها.

أود أن أشكر طلابي وزملائي في جامعة أدنبره على تعليقاتهم وتفاعلهم مع محاضراتي التي شكلت أساساً لهذا الكتاب. وكذلك الشكر لجميع المراجعين الذين قدموا التعليقات المفيدة، الكثير منهم مجهول، لكن هناك من قدم مقترحات أكثر قيمة ضُمنت في الكتاب أمثال: دينس يون (Denise Young)، كرايق هينيك (Craig Heinicke)، جون هتون (John Hutton)، كاي لي (Kai Li) وجين سوبر (Jean Soper).

وكذلك أعبرى عن امتناني بصفة خاصة إلى ستيف هاردمان (Steve Hardman) من مؤسسة جون وايلي (John Wiley) لتشجيعه ونصائحه المفيدة في مجال النشر التى ظل يقدمها طيلة فترة إعداد الكتاب. كما أود أيضا أن أعبر عن تقديري العميق لزوجتي ليزا (Lise)، لدعمها وتشجيعها لي خلال فترة إعداد هذا الكتاب.

مقدمة الطبعة الثانية

عندما كنت أعد الطبعة الجديدة من كتابي هذا، حاولت الأخذ في الاعتبار الملاحظات التي وردت من زملائي الذين قاموا بتدريس الطبعة الأولى والمراجعين الذين طلب منهم الناشر (Wiley) تقييم خطتي للطبعة الجديدة، إضافة إلى خبرتي الشخصية. فيما يتعلق بخبرتي الشخصية، فلقد قمت بتدريس الطبعة الأولى في ثلاث جامعات مختلفة هي: أدنبره، جلاسقو، ليسيستر (Edinburgh, Glasgow & Leicester) وعلى ثلاثة مستويات مختلفة. فقد قمت بتدريسه في مقرر السنة الثالثة (وذلك للطلاب غير المتخصصين في الاقتصاد ولديهم خلفية قليلة أو ليس لديهم خلفية في الإحصاء)، وفي مقرر السنة الثانية (وذلك للطلاب الذين لديهم معرفة متوسطة في الاقتصاد، مع وجود معرفة قليلة أو عدم وجودها في الإحصاء)، وفي مقرر السنة الأولي (وذلك للطلاب الذين يتلقون تحليل البيانات الاقتصادية لأول مرة). واستناداً إلى آراء الطلاب ومشاركتهم، فإن هذا الكتاب يمكن تدريسه بنجاح في كل هذه المستويات الأكاديمية. وأخبرني زملائي الأساتذة بأن الكتاب استخدم بنجاح لتدريس طلاب إدارة الأعمال وطلاب ماجستير إدارة الأعمال.

في هذه الطبعة من الكتاب لم يلغَ أي شيء من الطبعة الأولى (ماعدا التصحيحات المحدودة، أو تصحيح الأخطاء الطباعية والتغييرات الصياغية). إلا أنه تم إضافة الكثير من المواد والعناصر الجديدة، وذلك بغية تقديم تفاصيل حول أدنى خلفية رياضية مطلوبة لدراسة الكتاب. وبعضها يقدم توضيحاً أكثر للمفاهيم الأساسية مثل الأرقام القياسية، وبعضها يقدم وصفاً أكثر لمصادر البيانات. وفي كافة أجزاء هذه الطبعة، حرصت على أن تكون المادة أكثر وضوحاً، بحيث يصبح في الإمكان فهم المفاهيم المتعلقة بتحليل البيانات الاقتصادية بسهولة. وفي

ضوء تدريس الكتاب في مقررات إدارة الأعمال، فقد أضفت مادة ذات علاقة بطلاب إدارة الأعمال، وبخاصة الذين يدرسون مادة التمويل (Finance).

ومازلت متمسكاً بكل التعليقات التي أوردتها في مقدمة الطبعة الأولى لهذا الكتاب، وبخاصة تلك التي تعبر عن امتناني لكل الذين ساعدوني بتعليقاتهم ومقترحاتهم البناءة. وأود أن أضيف لقائمة من قمت بشكرهم في تلك المقدمة أسماء كل من جوليا داربي (Julia Darby)، كريستيان سكريدي قليدتش (Kristian Skrede Gleditsch) وهيلاري لاميسون (Hilary Lamaison) وكل طلابي من جامعات أدنبره (Edinburgh)، جلاسقو (Glasgow)، ليسيستر (Leicester).

الفصـــــل الأول

مقدمــة

(Introduction)

هناك عدة فئات من الاقتصاديين المتخصصين الذين يعملون في العالم اليوم. فالاقتصاديون الأكاديميون العاملون في الجامعات يقومون باشتقاق واختبار النماذج النظرية لمختلف جوانب الاقتصاد، والاقتصاديون العاملون في الخدمة المدنية يقومون بدراسة ما حققته وما لم تحققه السياسات الحكومية التي يجري العمل بها، أما الاقتصاديون العاملون في البنوك المركزية فيقدمون المشورة حول ما إذا كان يجب تغيير سعر الفائدة، في حين يقوم الاقتصاديون العاملون في القطاع الخاص بتقديم توقعاتهم المستقبلية للمتغيرات الاقتصادية، مثل تحركات سعر الصرف، وأثر تلك المتغيرات في صادرات ذلك الاقتصاد.

تُعد القدرة على معالجة البيانات مهارة مهمة للغاية لكل هؤلاء الاقتصاديين، إذ تمكنهم هذه المهارة من الفصل بين النظريات المتداخلة، وتقديم توقعات متعلقة بأثر التغيرات في السياسات، أو التوقع (التنبؤ) بما يمكن أن يحدث في المستقبل، حيث من الضروري البحث عن الحقائق، كما أننا نجد في علم الاقتصاد، وتحت تصرفنا، حقائق كثيرة (في شكل بيانات) يمكننا تحليلها بطرق مختلفة لإلقاء الضوء على قضايا اقتصادية عديدة.

إن الهدف من هذا الكتاب هو تقديم أساسيات تحليل البيانات بطريقة مبسطة غير رياضية، وذلك من خلال تأكيد أهمية الرسوم البيانية والتفسير اللفظي (الكلامي)، وتركز هذه الطريقة على الأدوات التي يستخدمها الاقتصادي في الحياة العملية (أهمها الانحدار Regression)، ويطور مهارات الحاسب الآلي التي تُعد ضرورة أساسية في أي مجال من مجالات العمل التي يرغب طالب الاقتصاد في الالتحاق بها.

ولتقديم شرح إضافي لما يقدمه هذا الكتاب، فإنه من الأفضل البدء في مناقشة مفهوم الاقتصاد القياسي. فيطلق "الاقتصاد القياسي" على الدراسة التي تستخدم أدوات وأساليب كمية لتحليل البيانات الاقتصادية، حيث إن مجال الاقتصاد القياسي

يعتمد على نظرية الاحتمالات والنظريات الإحصائية، وهو مجال رياضي بحت. في حين نجد أن هذا الكتاب لا يسعى إلى شرح النظرية الإحصائية ونظرية الاحتمالات، كما أنه لا يحتوي على صيغ رياضية كثيرة، ومن هذين المنطلقين فإن هذا الكتاب يختلف عن كتب الاقتصاد القياسي التقليدية، إذ يهدف إلى تدريس معظم الأدوات العملية التي يستخدمها حالياً الاقتصاديون القياسيون التطبيقيون.

إن الكتب التي تقوم بتعليم الطالب عن المفاتيح التي يجب أن يضغط عليها في لوحة مفاتيح الحاسب الآلي فقط دون إفهامه ما يقوم به جهاز الحاسب الآلي عند قيامه بعملية الضغط، تعرف بشكل عام على أنها كتب تمزج الأشياء ببعضها "كتب طبخ". وهذا الكتاب ليس واحداً من تلك الكتب. وقد يعترض بعض الاقتصاديين القياسيين على هذه النقطة بالقول: "كيف لكتاب يعلم الطالب كيفية استخدام أدوات القياس الاقتصادي دون تعليمه أساسيات نظرية الاحتمالات والإحصاء؟"، والإجابة عن ذلك هي أن الكثير مما يقوم به الاقتصاد القياسي في الحياة العملية يمكن فهمه بطريقة تلقائية دون الحاجة للرجوع إلى نظرية الاحتمالات أو الإحصاء. إن مضمون هذا الكتاب يتمثل في أن معظم الأدوات التي يستخدمها الاقتصاديون القياسيون يمكن التعامل معها من خلال فهم عميق لمفهوم الارتباط (Correlation) وتعميماته، والانحدار (Regression). فإذا فهم الطالب الارتباط والانحدار فهماً جيداً فإنه يمكنه فهم معظم ما يقوم به الاقتصاديون القياسيون. في أغلب الحالات يمكن القول إن الانحدار يمكن أن يكشف معظم المعلومات في مجموعة البيانات. يضاف إلى ذلك أن كلاً من الارتباط والانحدار هما مفهومان بسيطان إلى حد كبير، ويمكن فهمهما بطريقتين؛ الأولى تعتمد على استخدام التفسير الشفهي (الكلامي)، والثانية تعتمد على استخدام الرسوم البيانية. وتمثل هاتان الطريقتان الأساس لتوضيح المفاهيم الأكثر صعوبة، كما يمكن استخدامهما لتحليل العديد من أنواع البيانات الاقتصادية.

يركز هذا الكتاب على تحليل البيانات الاقتصادية، ولا يعنى بجمع البيانات الاقتصادية مع وجود بعض الاستثناءات، فهو كتاب يتعامل مع البيانات كما هي معطاة، ولا يوضح الكيفية التي تم بها جمع هذه البيانات. على سبيل المثال، لا يوضح هذا الكتاب كيف تم جمع الحسابات القومية أو كيف يتم تصميم وإجراء مسوحات سوق العمل، إنما ببساطة يعلم القارئ كيف يستفيد من البيانات التي تم جمعها.

تنطلق النظرية الإحصائية دائماً من التعريف الأساسي للمفاهيم العامة، وما يتبعه من نقاش حول كيف أن هذه المفاهيم وثيقة الصلة بأمثلة معينة. أما هذا الكتاب فيحاول القيام بعكس ذلك، أي أنه يحاول بحث المفاهيم العامة من خلال أمثلة معينة. وفي بعض الحالات لا يتم تقديم تعاريف أساسية، على سبيل المثال تعتبر قيم الاحتمالات (P-values) ودرجات الثقة مفاهيم إحصائية مهمة، وكذلك يقدم معايير تتعلق بدقة خط الانحدار (انظر الفصل الخامس). يستخدم الفصل المشار إليه الأمثلة والرسوم البيانية والتفسير الشفهي (الكلامي) لإيضاح كيفية استخدام هذه الأمثلة على أرض الواقع. إلا أنه لا يتم إعطاء تعريف أساسي لأي قيم احتمالات (P-values) أو اشتقاق لدرجة ثقة معنوية، لأن هذا يتطلب دراسة نظريات الاحتمالات والإحصاء، ولا يعتبر ذلك ضرورياً لاستخدام هذه الأساليب بشكل ملموس في الواقع العملي. أما عن القارئ الذي يرغب في معرفة المزيد عن النظرية الإحصائية التى تُعد منطلقاً لهذه الأدوات - فهناك العديد من الكتب المتاحة في هذا المجال مثل:

Introductory Statistics for Business and Economics by Thomas Wannacott and Roland Wannacott (Fourth Edition John Wiley & Sons, 1995).

أما لهؤلاء الذين يهتمون بالكيفية التي يتم بواسطتها تطبيق النظرية الإحصائية في بناء نماذج الاقتصاد القياسي فهناك كتب مثل: Undergraduate Econometrics by R.Carter Hill, William E.Griffiths and George G.Judge (Second ed.,Jown Wiley & Sons,2000).

إن هذا الكتاب يعكس إيمان المؤلف بأن استخدام الأمثلة المركزة المتنوعة يعد أفضل طريقة لتعلم تحليل البيانات؛ لذلك اشتمل كل فصل على العديد من الأمثلة لإيضاح المفاهيم الرئيسة. وهناك مأخذ على هذه الطريقة يتمثل في أن بعض الطلاب قد يفهمون أن وجود العديد من الأمثلة يعني أن هناك أعداداً لا حصر لها من المفاهيم التي يجب عليهم التمكن منها قبل أن يصبحوا جاهزين لممارسة الاقتصاد القياسي، وذلك لا يمثل الواقع، لأن المفاهيم الأساسية الموجودة بهذا الكتاب ليست كثيرة، إلا أنها تظهر بصورة متكررة في صورة مشكلات ومجموعات بيانات مختلفة ومتنوعة. بمعني آخر فإن الأسلوب الأمثل لتعليم مبادئ الاقتصاد القياسي يتمثل في توضيح مفاهيمه الخاصة وتكرارها بأنماط وتطبيقات مختلفة.

تنظيم الكتاب:

لقد حاولت خلال تنظيم هذا الكتاب، التقيد بنفس النهج العام الذي ذكرته سابقاً. إذ يغطي كل فصل من فصوله موضوع معين ويشتمل على نقاش عام حوله. إلا أن معظم الفصول قد خصصت للأمثلة القياسية التي توضح وتقدم، في بعض الحالات، مفاهيم أساسية مهمة.

كذلك يشتمل كل فصل على التمارين التي تشرح بشكل أعمق تلك المفاهيم. أما البيانات المطلوبة للاستخدام في الأمثلة القياسية والتمارين فيمكن الحصول عليها من خلال الموقع الإلكتروني للكتاب وهو:

http://www.wileyeurope.com/go/koopdata2ed

ومن خلال إضافة بيانات حقيقية واقعية نأمل، ليس وحسب، بأن يقوم الطلاب بإعادة تطبيق الأمثلة، بل أن يشعروا باليسر في توسيع أو اختبار البيانات أو كليهما بمختلف الأساليب.

فمن الضروري استخدام بيانات حقيقية واقعية إذا أراد الطلاب إتقان المادة النظرية أو تطبيق الأساليب الموضحة في هذا الكتاب.

صممت الأمثلة القياسية (التطبيقية) التي اشتمل عليها الكتاب للاستخدام مع برنامج إكسل (Excel). كما أن الموقع الإلكتروني للكتاب يحتوي على ملفات إكسل. برنامج إكسل يتكون من برامج جاهزة دارجة الاستخدام، وهو أيضا أحد البرامج التي يستخدمها طلاب الاقتصاد في ممارستهم العملية. وعلى كل، فإنه يمكن تحليل البيانات باستخدام العديد من البرامج الجاهزة، وليس الأمر مقصوراً على برنامج إكسل. إن العديد من البرامج يمكنها التعرف على ملفات إكسل، ومن ثم يمكن لهذه البرامج أخذ البيانات مباشرة من هذه الملفات. إضافة إلى ذلك، فإن الموقع الإلكتروني للكتاب يحتوي على ملفات جميع البيانات في شكل ملفات نصوص معيارية ASCII text. كما أن الملحق (ب) في آخر الكتاب يقدم تفاصيل أكثر حول هذه البيانات.

يُلاحظ في هذا الكتاب أن المادة الرياضية وضعت في أضيق نطاق. إلا أن في بعض الحالات القليلة نجد الرياضيات قدمت بتوسع أعمق، أما بالنسبة للطلاب الذين لديهم خلفية رياضية فيمكنهم الاطلاع على الملاحق الموجودة في آخر بعض الفصول، علماً أنه يمكن للطلاب تجاهل هذه المادة من غير أن يتأثر استيعابهم للمفاهيم الأساسية.

قسمت محتويات الكتاب بطريقة منطقية إلى قسمين. فالفصول من (1 إلى 7) تغطي كل المادة الأساسية المتعلقة بالتعبير بيانياً عن أساليب الارتباط والانحدار. وتكفي دورة دراسية قصيرة لتغطية هذه المادة. أما الفصول (8 إلى 12) فتركز على مواضيع السلاسل الزمنية كما تقوم بتحليل بعض النماذج القياسية المتقدمة المستخدمة في الوقت الحاضر. إن التركيز على الفهم البديهي للانحدار يعني أن هذه المادة يجب أن تكون سهلة في متناول الطلاب. ويجدر التنويه إلى أن الطلاب قد يجدون الفصول الأخيرة (8 إلى 12) أكثر صعوبة من الفصول (1 إلى 7).

خلفية مفيدة:

كما سبق ذكره، يفترض هذا الكتاب توافر خلفية رياضية بسيطة لدى الطالب قبل المرحلة الجامعية وخاصة الخلفيات التالية:

1- المعرفة بالمعادلات الرياضية البسيطة. على سبيل المثال، تم استخدام معادلة الخط المستقيم بصورة متكررة في هذا الكتاب.

2- المعرفة بأساليب الرسم البياني البسيط. على سبيل المثال، هذا الكتاب مليء بالرسوم البيانية التي ترسم مستويات متغير في مواجهة متغير آخر (أو ما يعرف بأشكال الانتشار XY-plot).

3- المعرفة بعمليات الجمع ($\sum$ Summation Operator) تكون مفيدة في بعض الأحيان.

4- في بعض الحالات القليلة هناك استخدام للوغاريثمات.

أما فيما يتعلق بالقارئ الذي ليس لديه خلفية عن هذه المواضيع فإن هناك ملحقاً في آخر هذا الفصل يقدم مقدمة موجزة عنها. بالإضافة إلى ذلك، فقد تمت مناقشة هذه المواضيع في العديد من كتب الرياضيات الأساسية الأولية.

هذا الكتاب يركز على استخدام الحاسب الآلي بشكل كبير، حيث تم فيه شرح الكثير من الموضوعات المتعلقة بالحاسب الآلي. هناك الكثير من برامج الحاسب الآلي الجاهزة التي يمكن استخدامها لتنفيذ الخطوات المشروحة في هذا الكتاب. في المواضع التي أكتب فيها عن برامج الحاسب الآلي، سوف استخدم لغة الجداول الإلكترونية (Spreadsheet) أكثر من استخدام برامج إحصاء متخصصة أو برامج اقتصاد قياسي مثل E-views, Stata or Micro Fit [1]. فأنا أتوقع أن الطالب يعرف أساسيات برنامج Excel (أو حتى أي نوع من برامج الحاسب الآلي التي يستخدمها / أو تستخدمها). بمعنى آخر، أنه يجب على الطلاب فهم أساسيات مصطلح الجداول الإلكترونية (Spreadsheet)، بحيث يكونون قادرين على فتح إعدادات البيانات، قص، نسخ ولصق البيانات... الخ.

فإذا كانت هذه المادة غير واضحة للطالب فإنه توجد تعليمات مبسطة في تعليمات المساعدة ببرنامج إكسل (تحت قائمة Help).

أما عن التعمق في استخدام الحاسب الآلي (ولهؤلاء الذين يريدون بسهولة معرفة الكثير عن جانب الحوسبة الآلية في تحليل البيانات) فقد يكون من المناسب لهم الاطلاع على كتاب Computing Skills for Economists الذي كتبه Guy Judge (John Wiley & Sons, 2000) فهو يعد بداية صائبة لهم.

ملحق رقم (1-1) المفاهيم الرياضية المستخدمة في الكتاب:

يستخدم هذا الكتاب أدنى قدر من الرياضيات، معتمداً بدلاً عن ذلك على الحدس البديهي والرسوم البيانية لتطوير استيعاب المفاهيم الأساسية (بما في ذلك فهم كيفية تفسير الأرقام التي تنتجها برامج الحاسب الآلي مثل برنامج إكسل). إن دراسة مادة الرياضيات في مرحلة ما قبل الجامعة تكفي لأن تعطي معظم الطلاب

الخلفية المعرفية التي يحتاجون إليها. ومع ذلك، هناك قائمة بالمفاهيم المستخدمة في هذا الكتاب مع توصيف مختصر لها.

معادلة الخط المستقيم:

يهتم الاقتصاديون دائما بالعلاقة بين متغيرين، (أو أكثر). الأمثلة على هذه المتغيرات تشمل أسعار المنازل، إجمالي الناتج المحلي (GDP)، أسعار الفائدة، ... الخ. في كتابنا هذا فإن المتغير هو الشيء الذي يهتم به الاقتصادي ويمكنه جمع بيانات عنه.

استخدمت في هذا الكتاب الحروف الكبيرة (مثل Y أو X) لتعريف المتغيرات. الطريقة الأكثر انتشاراً لتعريف العلاقة هي من خلال مفهوم الدالة. التعريف الرياضي العام للدالة بالنسبة للمتغير X هو $f(x)$ لذلك وعلى سبيل المثال، إذا كان الاقتصادي مهتماً بالعوامل التي تفسر أو تشرح لماذا تكون بعض المنازل أكثر قيمة من المنازل الأخرى، فقد يعتقد أن سعر المنزل يعتمد على حجمه. رياضياً، فقد يجعل Y تعرف المتغير "سعر المنزل" و X تعرف المتغير "حجم المنزل"، وحقيقة أن قيمة Y تعتمد على X تكتب رياضياً على النحو التالي:

$$Y = f(x)$$

هذا النظام يجب قراءته "Y دالة في X" ويتضمن الفكرة بأن قيمة Y تعتمد على قيمة X. هناك العديد من الدوال التي يمكن للفرد أن يستخدمها، إلا أنني سوف أركز في هذا الكتاب على الدوال الخطية. لذلك لن استخدم أسلوب الدوال "$f(x)$" في هذا الكتاب.

معادلة الخط المستقيم (التي تم تسميتها أعلاه بالدالة الخطية) تم استخدامها في كل هذا الكتاب. أي خط مستقيم يمكن كتابته في شكل المعادلة التالية:

$$Y = \alpha + \beta X$$

حيث α وβ هما معاملات، تحدد خطاً معيناً. لذلك وعلى سبيل المثال، عند وضع $\alpha = 1, \beta = 2$ فإنها تعرف خطاً واحداً معيناً في حين أن $\alpha = 4, \beta = -5$ تعرف خطا مختلفاً.

من الممكن أن يكون سهلاً فهم الخطوط المستقيمة من خلال استخدام الرسم البياني (وقد يكون من المفيد بالنسبة لك رسم إحداها في هذه المرحلة من الكتاب). وبالنسبة للرسم البياني Y X (بمعنى آخر، يقيس أحدهما قيم Y على المحور الرأسي وقيم X على المحور الأفقي) فإن أي خط يمكن تعريفه من خلال تقاطعه وميله. التقاطع هو قيمة Y عندما تكون X=0 (بمعنى آخر، النقطة التي يقطع فيها الخط محور Y). الميل هو قياس كم التغير في Y عندما تتغير X. وعادة، هو حجم التغير في Y عندما تتغير X بوحدة واحدة. أما عن الطالب الذي له إلمام بالتفاضل والتكامل، فإن الميل هو المشتقة الأولى (First Derivative) $\frac{dY}{dx}$.

رموز الجمع:

تُستخدم الرموز السفلية لبعض المتغيرات في أجزاء من هذا الكتاب للإشارة إلى المشاهدات المختلفة وتمييزها عن المتغير. على سبيل المثال، قد يكون الاقتصادي المتخصص في مجال اقتصاديات العمل مهتما بأجر كل فرد من المائة شخص الذين تضمهم صناعة معينة. فإذا استخدم الاقتصادي الحرف Y لتعريف هذا المتغير، بعد ذلك سوف يحصل على قيمة Y للفرد الأول، قيمة Y للفرد الثاني ... الخ. النظام الموحد لهذه الحالة هو استخدام الترقيم بحيث إن Y_1 هي أجر الفرد الأول، Y_2 هي أجر الفرد الثاني ... الخ.

في بعض السياقات، قد يكون من المفيد الحديث عن فرد يختص بصفة معينة وتعريفه بـ $i-th$ حيث يمكننا كتابة Y_i لأن i =1,...,100 للتعريف بمجموعة الأجور بالنسبة لكل الأفراد.

وبعد إنشاء نظام الترقيم، يمكن الآن تقديم مفهوم معامل الجمع $\sum$. ففي العديد من الحالات نريد أن نجمع المشاهدات أو البيانات المشاهدة (على سبيل المثال، عند حسابك للمتوسط فإنك تقوم بجمع كل المشاهدات والقسمة على عددها). الرمز اليوناني، $\sum$، هو معامل مجموع أو "جمع" والأرقام إلى الأعلى وإلى الأسفل من الرمز $\sum$ تشير إلى عدد المشاهدات التي تم جمعها.

لذلك، وعلى سبيل المثال

$$\sum_{i=1}^{100} Y_i = Y_1 + Y_2 + \ldots + Y_{100}$$

وهذه المعادلة تجمع كل الأجور بالنسبة للأفراد المائة. وكأمثلة أخرى،

$$\sum_{i=1}^{3} Y_i$$

وهى تجمع أجور الأفراد الثلاثة الأوائل

$$\sum_{i=47}^{48} Y_i$$

تجمع أجور الأفراد رقمي (47 و 48) في الترتيب العام للأفراد.

في بعض الأحيان، يتم إلغاء الأرقام العلوية والسفلية عندما يكون ذلك واضحاً من خلال السياق (وذلك عندما يتم جمع كل الأفراد أو العناصر) ويكتب ببساطة:

$$\sum Y_i$$

اللوغاريثمات:

نظراً للعديد من الأسباب (والتي سوف يتم إيضاحها لاحقاً)، فإنه أحياناً لا يعمل الباحث مباشرة مع المتغير وإنما مع نسخة معدلة أو محولة من هذا المتغير.

العديد من هذه التعديلات أو التحويلات تكون بصورة مباشرة. فمثلاً، عند مقارنة دخول دول مختلفة، فإنه يتم استخدام متغير متوسط دخل الفرد من إجمالي الناتج المحلي (GDP). وهو نسخة معدلة من متغير إجمالي الناتج المحلي. ويمكن الحصول عليه من خلال قسمة إجمالي الناتج المحلي على عدد السكان.

هناك تعديل أو تحويل شائع بصفة خاصة، وهو التعديل اللوغاريثمي. اللوغاريثم (للقاعدة B) للرقم A، هو القوة التي يجب رفع B بها لتعطي الرقم A. ويرمز لذلك بـ: $Log_B(A)$، لذلك وعلى سبيل المثال، إذا كانت A=100,B=10 فإن اللوغاريثم يساوي 2 ونكتب $Log_{10}(100) = 2$. وهي كذلك ما دام $10^2 = 100$.

من الشائع في علم الاقتصاد العمل بما يسمي اللوغاريثم الطبيعي والذي يرمز له بـ $B = e$ حيث $e \approx 2.71828$. ولن نقوم بتوضيح من أين جاءت e أو لماذا تم اختيار هذه القاعدة غير مألوفة الشكل. ويرمز لمعامل اللوغاريثم الطبيعي بـ ln، بمعني آخر $ln(A) = Log_e(A)$.

في هذا الكتاب لا يتطلب الأمر فهم المادة العلمية في الفقرة السابقة. الأمر الرئيس الذي يجب تذكره هو أن معامل اللوغاريثم الطبيعي هو معامل شائع (وذلك لأسباب سوف يتم إيضاحها لاحقاً) ويتم تعريفه بـ ln(A). في الواقع العملي، يمكن حسابه بسهولة في برامج مثل إكسل أو باستخدام آلة حاسبة.

ملاحظـــات ختامية:

1- أتوقع أن معظم قراء هذا الكتاب يتعاملون مع برنامج إكسل Excel (أو جدول إلكتروني Spreadsheet مشابه أو برامج إحصائية أخرى) من خلال معامل الحاسب الآلي في جامعاتهم أو على حاسباتهم الآلية في منازلهم (مع ملاحظة أن بعض المنهجيات المستخدمة في هذا الكتاب تتطلب تحليلات إكسل الإضافية Excel Analysis Toolpak add-in) والتي قد لاتكون مضمنة في بعض التركيبات الأساسية لحزمة برامج أوفيس (Microsoft Office).

ربما تكون برامج الحاسب الآلي مكلفة، خاصة للطالب الذي لايجيد التعامل مع إكسل ولديه صعوبات مالية. لذا عليه الرجوع إلى البرامج الإحصائية المجانية التي صممت باستخدام برامج المصدر المفتوح. وتعتبر حزمة برامج R.Zelig مثالاً جيداً لهذه المجموعة وهى متاحة على الموقع التالي على شبكة الإنترنت http://gking.harvard.edu/zelig/.

الفصل الثاني

أساسيات التعامل مع البيانات

(Basic Data Handling)

يقدم هذا الفصل أساسيات التعامل مع البيانات الاقتصادية. و يركز على أربعة مجالات مهمة وهي:

1- تنوع البيانات التي غالباً ما يستخدمها الاقتصاديون.

2- المصادر التي يحصل منها المحلل الاقتصادي على البيانات [1].

3- شرح لأنواع الرسوم البيانية التي تستخدم على نطاق واسع لعرض المعلومات التي تتضمنها البيانات.

4- عرض للمقاييس الحسابية البسيطة، أو الإحصاءات الوصفية، التي غالباً ما تقدم لتلخيص الجوانب الرئيسية لمجموعة البيانات.

أنواع البيانات الاقتصادية:

يقدم هذا القسم الأنواع الشائعة للبيانات ويُعرف المصطلحات المتعلقة باستخدامها.

بيانات السلاسل الزمنية (Time Series Data):

تقيس بيانات الاقتصاد الكلي ظاهرة معينة مثل إجمالي الناتج المحلي الحقيقي (المعروف بـ GDP)، أسعار الفائدة، عرض النقود...الخ. هذه البيانات يتم جمعها في أوقات معينة (على سبيل المثال، سنوياً). وعلى الجانب الآخر، تقيس البيانات المالية ظاهرة معينة مثل التغيرات في أسعار الأسهم. هذا النوع من البيانات يتم جمعه بصورة أكثر تكراراً من سابقتها، على سبيل المثال، يومياً أو حتى كل ساعة. وفي كل هذه الأمثلة، تكون البيانات مربوطة بالوقت وتسمى بيانات السلاسل الزمنية (Time Series). الظاهرة التي نقوم بقياسها (على سبيل المثال إجمالي الناتج المحلي أو الأجور أو أسعار الفائدة ... الخ) تسمى بالمتغير. بيانات السلسلة الزمنية يمكن ملاحظتها أو مشاهدتها من خلال تكرارات متعددة.

والتكرارات شائعة الاستخدام هي: سنوياً (يمكن ملاحظة المتغير كل سنة) ربع سنوي (أربع مرات في السنة)، شهرياً، أسبوعياً، أو يومياً.

في هذا الكتاب، سوف نستخدم الحرف Y_t للإشارة إلى إحدى مشاهدات المتغير Y (على سبيل المثال إجمالي الناتج المحلي الحقيقي) في الزمن أو الفترة t. وعندما تتراوح سلسلة البيانات بين الفترة t=1 و t=T ، فإننا نستخدم "T" للإشارة إلى الرقم الإجمالي للفترات الزمنية التى تغطيها البيانات. ولإعطاء مثل على ذلك، إذا أردنا استخدام بيانات إجمالي الناتج المحلي الحقيقي السنوية من 1946م إلى 1998م – وهي فترة 53 سنة - فإن t=1 تشير إلى سنة 1946م، t=53 تشير إلى سنة 1998م و T=53 العدد الإجمالي للسنوات. وبذلك تكون Y_1 هي إجمالي الناتج المحلي الحقيقي في عام 1946م، Y_2 هي إجمالي الناتج المحلي الحقيقي للعام 1947م ... الخ. وجرت العادة على عرض بيانات السلاسل الزمنية وفقاً لتسلسلها الزمني.

يتطلب العمل مع بيانات السلاسل الزمنية غالباً بعض الأدوات الخاصة، والتي نوقشت في الفصول (8-11).

البيانات المقطعية (Cross-Sectional Data):

في مقابل ما ذكر أعلاه، غالباً ما يعمل المتخصصون في الاقتصاد الجزئي ببيانات تتميز بالوحدات المفردة. هذه الوحدات قد تكون للأفراد، للشركات أو الدول. المَثل الشائع لذلك يتمثل في البيانات المتعلقة بالعديد من مختلف الأفراد في المجموعة الواحدة، مثل أجور كل الأفراد في شركة معينة أو صناعة. ففي مثل هذه البيانات المقطعية فإن مسألة ترتيب البيانات لا تكون مهمة (بخلاف الأمر في بيانات السلاسل الزمنية).

في هذا الكتاب استخدمنا الحرف Y_i للإشارة إلى بيانات العنصر أو المفردة i بالنسبة للمتغير Y للفرد i. وتبدأ المشاهدات في مجموعة البيانات المقطعية من المفردة i=1 إلى N.

وكالعادة، فإن N تشير إلى عدد الوحدات / المفردات المقطعية (على سبيل المثال، عدد الأفراد الذين أجري المسح عليهم). فمثلاً إذا أراد اقتصادي متخصص في مجال اقتصادات العمل مسح N=1000 عامل في صناعة الحديد الصلب، سائلاً كل فرد منهم أسئلة مثل ما هو حجم إنتاجهم أو هل يتبعون لاتحاد عمالي؟ في هذه الحالة ، فإن Y_1 سوف تساوي الأجر (أو عضوية الاتحاد العمالي) المسجلة بواسطة العامل الأول، Y_2 الأجر (أو عضوية الاتحاد العمالي) المسجلة بواسطة العامل الثاني، وهكذا.

وعلى نفس المنوال، قد يسأل الاقتصادي المتخصص في الاقتصاد الجزئي 100 (N=100) مندوب من الشركات الصناعية حول أرقام أرباحهم في الشهر الأخير. في هذه الحالة، فإن Y_1 تساوي الربح المسجل بواسطة الشركة الأولى، Y_2 هو الربح المسجل بواسطة الشركة الثانية، وهكذا حتى Y_{100}، وهو الربح المسجل بواسطة الشركة رقم 100.

التفريق بين البيانات النوعية والكمية:

يمكننا استخدام مجموعات البيانات السابقة لتوضيح تمييز مهم بين أنواع البيانات. فبيانات المبيعات التي جمعها المتخصص في الاقتصاد الجزئي سوف يكون لها رقم يقابل كل مؤسسة تم مسحها (على سبيل المثال مبيعات الشهر الأخير في الشركة الأولى التي تم مسحها هي 20000 جنيه إسترليني). وهذه تسمى بيانات كمية.

بينما عندما يسأل الاقتصادي المتخصص في مجال اقتصاديات العمل عن هل ينتمي أو لا ينتمي المستخدم الذي أجري المسح عليه لعضوية اتحاد أم لا؟ فإنه يتلقى إجابة إما بنعم أو لا. هذه الإجابات تسمى بيانات نوعية. هذه البيانات كثيراً ما تبرز في الاقتصاد عندما يكون هناك خيارات (على سبيل المثال، خيار أن تشتري أولا تشتري منتجاً، أن تستخدم النقل العام أو السيارة الخاصة، أن تنضم أو لا تنضم للنادي).

يقوم الاقتصاديون في العادة بتحويل هذه الإجابات النوعية إلى شكل بيانات رقمية. على سبيل المثال، قد يضع الاقتصادي المتخصص في مجال العمل (نعم) تساوي واحداً (Yes=1) و(لا) تساوي صفراً (No=0).

وقياساً على ذلك فإن $Y_1 = 1$ تعنى أن الفرد الأول الذي أجري المسح عليه ينتمي إلى عضوية اتحاد، $Y_2 = 0$ تعنى أن الفرد الثاني لاينتمى. وعندما يكون من الممكن أن تأخذ المتغيرات قيم 0 أو 1 فقط فإنها تسمي متغيرات صورية (وهمية) (dummy or binary variables). التعامل مع هذه المتغيرات الصورية هو موضوع الفصل السابع.

بيانات السلاسل الزمنية المقطعية (Panel data):

تجمع بعض مجموعات البيانات كلاً من عنصري السلاسل الزمنية والمقطعية، وتسمى هذه البيانات ببيانات السلاسل الزمنية المقطعية. يستفيد الاقتصاديون الذين يعملون على قضايا متعلقة بالنمو غالباً من بيانات السلاسل الزمنية المقطعية. على سبيل المثال، تتوافر بيانات إجمالي الناتج المحلي للعديد من الدول ابتداء من عام 1950م وحتى الوقت الحالي. بيانات السلاسل الزمنية المقطعية عن Y=GDP بالنسبة لـ 12 دولة أوروبية تشتمل على قيمة إجمالي الناتج المحلي لكل دولة في عام 1950م (12 مشاهدة من البيانات N=12)،

يتبعها قيمة إجمالي الناتج المحلي لكل دولة في عام 1951م (وهى N=12 مشاهدة أخرى) وهكذا. وعلى مدى فترة مقدارها T من السنوات، سوف تكون هناك مشاهدات من البيانات تساوى T×N عن Y. وعلى العكس من ذلك غالباً ما يتعامل المتخصصون في اقتصاديات العمل مع حزمة بيانات كبيرة يتم إنشاؤها من خلال طرح العديد من الأسئلة مثل ما حجم إنتاجهم كل سنة وعلى مدى سنوات عديدة.

سوف نستخدم الرمز Y_{it} للإشارة إلى بيانات حول المتغير Y بالنسبة للمفردة / العنصر i عند الزمن t. في مثال النمو الاقتصادي، Y_{11} هي إجمالي الناتج المحلي في الدولة 1، في السنة1، Y_{12} هي إجمالي الناتج المحلي للدولة 1 في السنة 2 ... الخ. في مثال اقتصاديات العمل، Y_{11} هي أجر الفرد الأول الذي أجري المسح عليه في السنة الأولى، Y_{12} هي أجر الفرد الأول الذي أجري المسح عليه في السنة الثانية ... الخ.

تحويلات البيانات: المستويات في مقابل معدلات النمو:

في هذا الكتاب افترضنا بشكل أساسي أن البيانات المتعلقة بالمتغير Y والتي نرغب في الحصول عليها سوف تكون متاحة بشكل مباشر. إلا أنه في الواقع العملي قد يطلب منك أخذ بيانات أولية من مصدر واحد، ثم تحويلها بعد ذلك إلى شكل مختلف من أجل تحليلاتك القياسية (التطبيقية). على سبيل المثال، قد تأخذ بيانات سلسلة زمنية أولية على المتغيرات W= (إجمالي الإنفاق على الاستهلاك)، و X = (الإنفاق على الطعام)، وإنشاء متغير جديد: Y = الجزء من الإنفاق الذي يتم تخصيصه للطعام. وهنا يكون التحويل Y=X/W.

تعتمد طبيعة التحويلات المطلوبة على المشكلة مدار البحث، لذلك فإنه من الصعب تقديم أي توصيات عامة على تحويل البيانات. وقد جرى مناقشة بعض الحالات الخاصة في فصول لاحقة. من المهم هنا تقديم تحويل عام واحد شائع الاستخدام من قبل الاقتصاديين القياسيين مع بيانات السلسلة الزمنية.

ولفهم المنطق العام وراء هذا التحويل، افترض أن لدينا بيانات سنوية عن إجمالي الناتج المحلي الحقيقي للفترة 1950- 1998م (بمعنى آخر بيانات 49 سنة) يشار إليها بـ Y_t حيث t =1 إلى 49.

في العديد من البحوث والتقارير القياسية، قد يكون إجمالي الناتج المحلي الحقيقي هو المتغير الذي يستحوذ على الاهتمام الأساسي. وسوف نسمي هذه السلسلة بمستوى (Level) إجمالي الناتج المحلي الحقيقي. على الرغم من أن كثيرا يهتمون بالكيفية التي ينمو بها الاقتصاد عبر الزمن، أو يهتمون بنمو إجمالي الناتج المحلي الحقيقي.

الطريقة السهلة لقياس النمو هي أن تأخذ سلسلة بيانات إجمالي الناتج المحلي الحقيقي وحساب نسبة التغير لكل سنة. نسبة التغير في إجمالي الناتج المحلي الحقيقي بين الفترة t والفترة t+1 يتم حسابها وفقا للمعادلة التالية[2].

$$\%change = \frac{(Y_{t+1} - Y_t)}{Y_t} x100$$

إن نسبة التغير في إجمالي الناتج المحلي الحقيقي كثيراً ما تسمى النمو في إجمالي الناتج المحلي أو التغير فيه. سوف تتم مناقشة بيانات السلاسل الزمنية بتفصيل أكثر في الفصول (8 – 11). ويكفي أن تعرف في هذه المرحلة أننا سوف نفرق من وقت لآخر بين مستوى متغير ما ومعدل نموه، وأنه من الشائع تحويل بيانات مستوى المتغير إلى معدلات نمو.

الأرقام القياسيـة (Index Numbers):

يـأتي العديـد مـن المتغيـرات التـي يتعامـل معهـا الاقتـصاديون فـي شـكل أرقـام قياسية، ويقدم الملحق رقم (2-1) في نـهاية هذا الفصل مناقشة مفصلة عن ماهيـة هذه الأرقام وكيف يتم حسابها. وعلـى الـرغم من ذلك، إذا أردت اسـتخدام رقـم قياسي في عملك التطبيقي، فإنه ليس من الضروري أن تكون لـديك معرفـة أكيـدة عن كيفية حساب الأرقام القياسية، ويكفي توافر فهم بـديهي جيـد عـن كيفيـة التعبيـر عـن الـرقم القياسـي. ووفقـا لـذلك، فقد قدمت فـي هـذا الكتـاب مناقـشة بديهيـة عـن الأرقام القياسية.

افترض أنك مهتم بدراسة معدل التضخم لدولة ما، وهذا المعدل هو عبارة عن قياس الكيفيـة التـي تتغيـر بهـا الأسـعار علـى المـدى الزمنـي. ويبـرز الـسؤال حـول الكيفيـة التـي نقـيس بهـا "الأسـعار" فـي دولـة مـا. حيـث يمكـن قيـاس سـعر الـسلعة الواحـدة (علـى سـبيل المثـال، الحليـب، البرتقـال، الكهربـاء، موديـل معـين لـسيارة، زوج الأحذية، ... الـخ) إلا أن الاهتمـام لا يتركـز علـى الـسلع المفـردة، وإنمـا علـى مستوى السعر في الدولـة ككل. وهـذا المفهـوم الأخيـر عـادة مـا يـتم تعريفـه بـسعر "السلة" وهو يشتمل على أنواع من السلع التي قد يشتريها المستهلك العـادي. تـتم ملاحظة سعر هذه السلة على مدى فترات زمنيـة منتظمـة وذلـك مـن أجـل تحديـد الكيفية التـي تتغيـر بهـا الأسـعار فـي الدولـة ككـل. إلا أن سـعر الـسلة لايـتم تحديـده عادة بشكل مباشر من قِبل الجهة الحكومية التي جمعت تلك البيانـات. وعلـى كـل، إذا تم إبلاغك بأن سعر سلعة معينة (على سبيل المثـال، أن سـعر البرتقـال يـساوي 35 بنساً)، فإنك قد أخبرت بشيء ذي معنى، أمـا إذا تـم إخبـارك بـأن (سـعر الـسلة لسلع ممثلة) هو 10.45 جنيه إسترليني فإن هذه العبـارة ليـست ذات معنـى بـشكل كبير. ولتوضيح الرقم الأخيـر، فإنك تحتـاج أن تعـرف ماذا يوجد بالتحديد في السلة

وبأي كميات. وبوجود ملايين السلع التي تشترى وتباع في الاقتصاد الحديث، فإن هناك كماً هائلاً من المعلومات التي يجب إعطاؤها أو تقديمها.

وفي ضوء مثل هذه القضايا، فإن البيانات كثيراً ما تأتي في شكل أرقام قياسية. الأرقام القياسية يمكن حسابها بعدة طرق، ومناقشتها هنا سوف تبعدنا عن الاهتمام الرئيسي لهذا الفصل (انظر الملحق رقم (2-1) لمزيد من التفاصيل). إلا أنه من المهم معرفة النقاط التالية في هذه المرحلة.

أولاً: تأتي الأرقام القياسية على شكل بيانات سلاسل زمنية.

ثانياً: عادة ما يتم اختيار فترة زمنية واحدة لتكون سنة أساس ويتم اعتماد الرقم 100 لمستوى السعر في سنة الأساس (بعض الأرقام القياسية تعتمد قيمة سنة الأساس بالرقم 1 بدلاً عن الرقم 100).

ثالثاً: يتم قياس مستويات الأسعار في السنوات الأخرى بنسب مئوية مقارنة مع سنة الأساس.

من المفيد في هذا السياق تقديم مثال لهذه المسائل كما يلي:

افترض وجود رقم قياسي للأسعار لفترة أربع سنوات بالقيم التالية:

$Y_1 = 100, Y_2 = 106, Y_3 = 109, Y_4 = 111$ ويمكن تفسير هذه الأرقام على النحو التالي: السنة الأولى اختيرت باعتبارها سنة أساس، ووفقاً لذلك فإن $Y_1 = 100$. الأرقام الخاصة بالسنوات الأخرى هي جميعها مقارنة بسنة الأساس ويتيح ذلك عملية حسابية بسيطة لتوضيح كيف تغيرت الأسعار منذ سنة الأساس. على سبيل المثال، $Y_2 = 106$ تعني أن الأسعار قد ارتفعت من 100 إلى 106، أي أن هناك ارتفاعاً في الأسعار بنسبة 6% منذ السنة الأولى. كذلك يمكن

ملاحظة أن الأسعار قد ارتفعت بنسبة 9% من السنة الأولى إلى السنة الثالثة وبنسبة 11% من السنة الأولى إلى السنة الرابعة.

ومادام أن نسبة التغير في الأسعار هو ما يطلق عليه التضخم، فإن الأرقام القياسية للأسعار تتيح للفرد المطلع على البيانات أن يعرف بسهولة ما هو معدل التضخم. بمعنى آخر، يمكنك التفكير في رقم قياسي للأسعار كأسلوب لعرض بيانات الأسعار لكي تكون سهلة التفسير والفهم.

يُعد الرقم القياسي للأسعار مقياساً جيداً لقياس التغيرات في الأسعار على المدى الزمني، ولكن لايجب استخدامه للحديث عن مستوى الأسعار. على سبيل المثال، لايجب الأخذ بالرقم القياسي للأسعار كمؤشر على ما إذا كانت الأسعار عالية أو منخفضة. ويمكن توضيح هذا بمثال بسيط:

تجمع كل من الولايات المتحدة الأمريكية وكندا بيانات عن أسعار المستهلك. افترض أن البلدين قررا جعل 1988م لتكون سنة أساس لأرقام أسعارها القياسية. وهذا يعنى أن الرقم القياسي للأسعار في عام 1988م للبلدين سوف يكون 100، إلا أنه لايعني أن الأسعار متساوية في البلدين في عام 1988م. اختيار عام 1988م لتكون سنة أساس تم بطريقة عشوائية، فإذا قررت كندا فجأة تغيير سنة الأساس إلى 1987م، فإن الأرقام القياسية للأسعار في سنة 1988م لن تكون متساوية في البلدين. في هذه الحالة فإن الأرقام القياسية للأسعار لايمكن استخدامها لإطلاق عبارات مثل: "الأسعار عالية في كندا مقارنة بالأسعار في الولايات المتحدة الأمريكية". إلا أنه يمكن استخدامها لحساب معدلات التضخم. وهذا يتيح لنا أن نقول: "إن التضخم في كندا (بمعنى آخر التغيرات في الأسعار) أعلى من التضخم في الولايات المتحدة الأمريكية".

التمويل هو مجال آخر غالباً تبرز فيه أهمية الأرقام القياسية للأسعار لأن المعلومات عن أسعار سوق الأسهم غالباً ما يتم تقديمها في هذا الشكل. بمعنى

آخر، أن مؤشرات ومقاييس أنشطة سوق الأسهم مثل متوسط داوجونز الصناعي DJIA، مؤشر فايناتشال تايمز FTSE ومؤشر ستاندارد آند بورز500 S&P500 هى جميعها أرقام قياسية للأسعار.

في نقاشنا حول هذا الموضوع، ركزنا على بعض الأرقام القياسية للأسعار التى تُعد حقاً أكثر الأرقام القياسية شيوعاً. لاحظ أن هناك أنواعاً أخرى من الأرقام القياسية (على سبيل المثال الأرقام القياسية للكميات)، ويجب توضيحها بنفس طريقة الأرقام القياسية للأسعار. أى أنها يجب أن تُستخدم كأساس لقياس الكيفية التى تتغير بها الظاهرة بالقياس على سنة أساس معلومة.

هذا النقاش حول الأرقام القياسية للأسعار هو المكان المناسب للحديث عن تحويل آخر يتم استخدامه للتعامل مع آثار التضخم. وكمثال على ذلك، فإنه يمكن حساب أكثر المقاييس شيوعاً للإنتاج في الاقتصاد: إجمالى الناتج المحلي (GDP) الذي يمكن حسابه من خلال تجميع قيم كل السلع النهائية المنتجة في الاقتصاد. إلا أن مجرد النظر إلى الكيفية التى يتغير بها إجمالي الناتج المحلي على المدى الزمنى قد تكون مضللة في الأوقات التى يكون فيها التضخم عالياً. فإذا كان معدل التضخم عالياً، فإن أسعار السلع سوف ترتفع وبالتالي فإن قيمتها سوف ترتفع على مر الزمن، حتى إذا لم تزد كمية السلع المنتجة. وما دام أن إجمالي الناتج المحلي يقيس قيمة كل السلع، فإنه سوف يكون مرتفعاً في الأوقات التى يكون فيها التضخم عالياً حتى إذا لم يتغير حجم الإنتاج. وهذا يجعل الباحثين راغبين في تصحيح أثر التضخم والطريقة التي يتم بها ذلك هى قسمة إجمالي الناتج المحلي على الرقم القياسي للأسعار (في حالة إجمالي الناتج المحلي، فإن اسم الرقم القياسي للأسعار هو مخفض إجمالي الناتج المحلي GDP Deflator). إجمالي الناتج المحلي الذي يتم تحويله بهذه الطريقة يسمى إجمالي الناتج المحلي الحقيقي. هذا التفريق بين المتغيرات الحقيقية والاسمية يعتبر مهماً في العديد من

مجالات الاقتصاد. ومن الأمور الأساسية التي يجب أن نتذكرها هى أن أي متغير حقيقي هو عبارة عن المتغير الاسمي مقسوماً على متغير السعر (وهو عادة الرقم القياسي للأسعار) وأنه قد تم إزالة آثار التضخم من المتغيرات الحقيقية.

إن الحالة التي تود فيها تصحيح معدل النمو في حالة وجود تضخم تختلف قليلاً، ففي هذه الحالة، فإن عملية إنشاء متغير حقيقي تتضمن طرح التغير في الرقم القياسي للأسعار من المتغير الاسمي. لذلك، وعلى سبيل المثال، فإن أسعار الفائدة الحقيقية هى معدلات سعر فائدة اسمية مطروحاً منها معدل التضخم (حيث يُعرف معدل التضخم بأنه معدل التغير في الرقم القياسي للأسعار).

الحصول على البيانات:

كل البيانات التى تحتاجها لاستيعاب المفاهيم الأساسية وإجراء التحليلات البسيطة التى يتناولها هذا الكتاب، يمكن تحميلها من الموقع الإلكتروني على شبكة الإنترنت والذي سبق ذكره، على الرغم من أنك قد تحتاج مستقبلاً أن تجمع البيانات التى تريدها لمقالة علمية، أو بحث أكاديمي أو تقرير. تأتى البيانات الاقتصادية من مصادر عديدة مختلفة ومن الصعب تقديم مقترحات عامة على عملية جمع البيانات. وفيما يلي بعض النقاط الرئيسة التي يجب عليك معرفتها عن البيانات الشائعة الاستخدام وأماكن وجودها.

إن معظم بيانات الاقتصاد الكلي يمكن جمعها من خلال نظام الحسابات القومية، والتي تكون متاحة عن طريق المطبوعات الحكومية، وبصورة متزايدة، في شكل رقمي في المكتبات الجامعية والحكومية. أما بيانات الاقتصاد الجزئي فتجمع عادة من خلال المسوحات الميدانية للقطاع العائلي، العمالة وقطاع الأعمال، والتي غالباً ما تكون متاحة من نفس المصادر.

وقد أصبح من الشائع، وبصورة متزايدة، أن يحصل الاقتصاديون على البيانات التى يحتاجونها من شبكة الإنترنت، وهناك الآن العديد من المواقع المناسبة على شبكة الإنترنت التي يمكن تحميل البيانات منها. ويجب عليك أن تأخذ حذرك مسبقاً من أن شبكة الإنترنت تنمو وتتغير بسرعة، حيث إن المعلومات والعناوين التى تقدم هنا قد تتغير عناوينها ومحتوياتها خلال فترة زمنية قصيرة. وعلى ذلك، فإن هذا القسم قد تم تقديمه لإعطاء أمثلة فقط عن ما يمكن الحصول عليه من شبكة الإنترنت، وهو بذلك لايتضمن كل ما يمكن الحصول عليه. ولمزيد من التفاصيل عن ما هو متاح على شبكة الإنترنت وكيفية الوصول إليه، ننصح بالاطلاع على كتاب Computing Skills for Economist للمؤلف Guy Judge.

وقبل أن تبدأ في البحث، يجب عليك أن تعرف أيضاً أن بعض المواقع تتيح للمستخدمين الوصول إلى البيانات مجاناً، في حين تأخذ مواقع أخرى رسوماً على البيانات. وهناك العديد من المواقع تقدم بيانات مجانية للمستخدمين غير التجاريين (مثل الجامعات) ويتطلب ذلك أن تسجل قبل أن يتاح لك الوصول للبيانات.

في هذا السياق فإن الموقع الأمريكي الأكثر فائدة هو "Resources for Economists on the Internet" "موارد للاقتصاديين على شبكة الإنترنت" (http://rfe.wustl.edu/EconFaq.html). هذا الموقع يشتمل على كل أنواع المعلومات والموضوعات الاقتصادية المثيرة للاهتمام التي يجب عليك بذل الوقت لاستكشافها. هذا الموقع يقدم أيضاً روابط ذات صلة للعديد من مصادر البيانات المختلفة. وهناك موقع آخر به روابط مفيدة هو:

National Bureau of Economic Research (http://www.nber.org/).

وعلى هذا الموقع هناك مصدر بيانات مفيد هو Penn World Table (PWT) إذ يقدم بيانات اقتصادية كلية عن مائة دولة ولسنوات عديدة. وسوف نعود لهذا الموقع لاحقاً في هذا الفصل.

وفي المملكة المتحدة هناك موقع MIMAS (Manchester Information & Social Services) وهو مدخل مفيد للعديد من مجموعات البيانات (http://www.mimas.ac.uk) وهذا الموقع يتطلب حالياً عملية تسجيل قبل الدخول إليه.

جدير بالذكر أن البيانات على شبكة الإنترنت غالباً ماتكون في شكل قائمة مبسطة على الشاشة، ويمكنك أن تقوم بنسخ البيانات يدوياً ثم بعد ذلك طباعتها في برنامج إكسل Excel. إلا أنه يمكن توفير الوقت من خلال حفظ البيانات في ملف باستخدام File/save as أو تحديد Highlight البيانات، ونسخها كمقتطفات ومن ثم لصقها في برنامج إكسل Excel.

ولإعطائك لمحة عن أنواع البيانات المتوافرة على شبكة الإنترنت، وكيف تبدو مواقع شبكة الإنترنت، سوف نركز على موقعين (أمريكي وبريطاني) شائعي الاستخدام على شبكة الإنترنت.

مثال: "Resources for Economists on the Internet"

فإذا تابعت الرابط المعنون Data على موقع:

Resources for Economists on the Internet

فإن الصفحة التالية سوف تظهر على الشاشة

Data

MS Macro and Regional Data (Data for the US Economy and its Regions)

Other V.S. Data (other types of US Data)

World and Non-US Data (Data from around the World)

Finance and Financial Markets (Data from Financial Markets)

Journal Data and Program Archives (academic journal archives)

فإذا ضغطت على زر الفأرة (الماوس) على أي من هذه الروابط المكتوبة بالخط الصغير (etalics) سوف تحصل على قائمة بالعديد من روابط شبكة الإنترنت الإضافية التي يمكنك الاتصال بها، وهى تشتمل على مختلف أنواع البيانات.

مثال: MIMAS

المادة التالية أُخذت مباشرة من موقع MIMAS. وبالطبع، لايمكنك فهم ماذا تعني كل العناوين المذكورة أدناه. إلا أنه يمكن ذكر القليل من الاختصارات أدناه على النحو التالي: (وهو المصدر الرئيسي لبيانات حكومة المملكة المتحدة) ONS='Office of National Statistics' وهو يجمع البيانات من العديد من الدول بما فيها الدول النامية IMF=International Monetarry Fund and OECD= Organization for Economic Co-operation and Development وهو يجمع بيانات للدول الصناعية.

الإحصاء السكاني ومجموعات البيانات المتعلقة به:

مدخل معلومات الإحصاء السكاني

1991 Local Base and small Area statistics-download a registration pack.

مكان العمل الخاص وإحصاءات الهجرة

1991 Samples of Anonymized Records.
1981 Small Area Statitics.
1991 Census Digitized Boundary Data.
1981 Digitized Boundary Data.
Table 100.
Census Monitor county/District tables.
Topic Statistics.
Population Surface Models.
Estimating with Confidence Data.
ONS ward and district level classifications.
GB Profiler.

The Longitudinal Study.

1971/81 Change File.

Postcode to ED/OA Directories.

Central Postcode Directory POSTZON File.

ONS Vital Statistics for Wards.

المسوحات الحكومية والمسوحات المتواصلة الأخرى

General Household Survey.

General Household Survey.

Labour Force Survey.

Quarterly Labour Force Survey.

Family Expenditure Survey.

Farm Business Survey.

National Child Development Study.

British Household Panel Study.

Health Survey for England.

بنوك بيانات السلاسل الزمنية للاقتصاد الكلي

ONS Time Series Databank.

OECD Main Economic Indicators.

UNIDO Industrial Statistics 3 digit level of ISIC code.

UNIDO Industrial Statistics 4 digit level of ISIC code.

UNIDO Commodity Balance Statistics Database.

IMF International Financial Statistics.

IMF Direction of Trade Statistics.

IMF Government Finance Statistics Yearbook.

ملاحـظـة:

يستطيع المستخدمون المسجلون فقط الوصول إلى مجموعات البيانات المشار إليها أعلاه.

يجب التنبيه إلى أن العديد من مجموعات البيانات المشار إليها أعلاه يمكن الوصول إليها والاستفادة منها مجاناً. يضاف إلى ذلك، أن معظم المكتبات الجامعية أو مراكز الحاسب الآلي تشترك في مختلف قواعد البيانات التى يستخدمها الطالب. ونوصيك بمراجعة مكتبة جامعتك أو مركز الحاسب الآلي لمعرفة مجموعات البيانات التى يمكنك الوصول إليها. في مجال التمويل، هناك العديد من قواعد البيانات المتميزة عن أسعار الأسهم ومعلومات محاسبية لكل أنواع الشركات ولعدة سنوات. ولسوء الحظ، أن مثل هذه البيانات غالباً ما تكون مكلفة، لذلك يجب عليك التأكد ما إذا كانت لجامعتك اشتراك في قاعدة بيانات مالية. وهنالك اثنان من أشهر هذه القواعد هما:

Datastream by Thomson Financial

(http://www.data_stream.com)

and

Wharton Research Data Services

(http://wrds.wharton.upenn.edu/)

أما فيما يختص بالبيانات المجانية، فهناك خيار محدود للوصول للبيانات المالية من خلال مواقع شبكة الإنترنت المشهورة مثل Yahoo،

(http://finance.yahoo.com)

ويقدم أيضا بنك الاحتياطي الفيدرالي الأمريكي فرع سانت لويس (Federal Reserve Bank of St. Louis) قاعدة بيانات مجانية تشتمل على بيانات واسعة ومتنوعة، فيها بعض السلاسل الزمنية المالية

http://research.stlouisfed.org/fred2/

وهناك العديد من المجالات التخصصية متاحة بالمجان على شبكة الإنترنت، على سبيل المثال في مجال الاقتصاد الرياضي والإحصاء هناك العديد من قواعد البيانات المتميزة المتاحة مجاناً أو بسعر رمزي. وبالنسبة لرياضة البيسبول فهناك موقع http://www.baseball1.com وهو موقع مليء بمجموعة بيانات عن هذه الرياضة. الإحصاءات الموجودة في قسم الرياضة التابع للجمعية الإحصائية الأمريكية يتوافر فيه روابط لمجموعات بيانات للعديد من أنواع الرياضة http://www.amstat.org/sections/sis/sports.html.

النصيحة العامة في هذا المجال هي أن بذل بعض الوقت في البحث في الإنترنت غالباً ما يكون ذا فائدة كبيرة.

التعامل مع البيانات: الأساليب البيانية:

في حال حصولك على بيانات، فإنه من المهم تلخيصها. وبعد هذه الخطوة لن يحتاج المطلع على تلك البيانات إلى معرفة خلفياتها المكونة من أرقام كبيرة قد تصل إلى المئات أو عدد المشاهدات الكثيرة التى تحتويها مجموعة البيانات الأولية الأصلية. وفي الواقع، يمكنك التعامل مع مجال الاقتصاد القياسي باعتباره مجالاً مخصصاً بالكامل لتطوير أساليب للتعامل مع ومعالجة البيانات لعرضها بشكل ملخص ومبسط ومفيد. وتعتبر الرسوم البيانية والجداول وسائل مفيدة جدا لتقديم بياناتك.

وهناك عدة أنواع من الرسوم البيانية والأشكال (مثل رسم دائرة بيانية، pie chart، رسم أعمدة bar chart ... الخ) وهناك طريقة مفيدة لمعرفة الرسوم البيانية (Charts) تتمثل في الاطلاع على معالج الرسوم البيانية (Chartwizard) في برنامج إكسل. في هذا القسم، سوف نقوم بتوضيح بعض أنواع الرسوم البيانية شائعة الاستخدام.

ومادامت معظم البيانات الاقتصادية إما في شكل بيانات سلاسل زمنية أو في شكل بيانات مقطعية، فسوف نقدم باختصار أساليب بسيطة للعرض البياني لكلا النوعين من البيانات.

الرسوم البيانية لبيانات السلاسل الزمنية:

تم رسم بيانات السلاسل الزمنية الشهرية لسعر صرف الجنيه الإسترليني مقابل الدولار الأمريكي مقاسة بنس لكل دولار (Pence per Dollar) خلال الفترة من يناير 1947م إلى أكتوبر 1996م باستخدام الرسم البياني الخطى (Line chart) المتوافر في برنامج إكسل (Chartwizard) كما هو موضح في الشكل رقم (2-1) أدناه (هذه البيانات متوافرة في ملف البيانات إكسل EXRUK.xls).

مثل هذه الرسومات تسمى الرسوم البيانية لبيانات السلاسل الزمنية. مجموعة البيانات المشار إليها تشتمل على 598 مشاهدة، لذا فهي كثيرة جداً لعرضها على القارئ في شكل جدول بيانات. ومع ذلك، يمكن للقارئ الإلمام بالملامح الرئيسية للبيانات بسهولة، وذلك من خلال النظر إلى الرسم البياني. ويمكن للقارئ، على سبيل المثال ملاحظة تجربة حكومة المملكة المتحدة لتثبيت سعر الصرف حتى نهاية عام 1971م (عدا تخفيض قيمة عملتها في شهر سبتمبر 1949م وشهر نوفمبر 1967م) والانخفاض التدريجي لقيمة الجنيه الإسترليني الذي تم تعويمه في منتصف السبعينيات من القرن الماضي.

ملاحظة:

- تخفيض قيمة العملة devalvation، عادة ما يتم باتخاذ قرار من الحكومة بذلك وتأخذه الأسواق على أنه أمر معطى.
- انخفاض قيمة العملة depreciation، هو نتيجة مباشرة لتقييم عوامل السوق لقيمة هذه العملة ولا دخل للسلطات النقدية بهذا الانخفاض بشكل مباشر.

شكل رقم (2-1) رسم بياني للسلسلة الزمنية لسعر صرف الجنيه الإسترليني مقابل الدولار الأمريكي

تمرين رقم (2-1):

(أ) أعد رسم الشكل رقم (2-1)

(ب) يحتوي ملف INCOME.xls بيانات من اللوغاريثمات الطبيعية للدخل الشخصي والاستهلاك في الولايات المتحدة الأمريكية من الربع الأول لعام 1954م إلى الربع الثاني من عام 1994م. ارسم شكلاً بيانياً واحداً لبيانات سلسلة زمنية تشتمل على كل من هذين المتغيرين. (لاحظ أن التاريخ 1954Q1 يشير إلى الربع الأول (يناير، فبراير، مارس لعام 1954م).

(جـ) حول بيانات الدخل الشخصي اللوغاريثمي إلى معدلات نمو. لاحظ أن نسبة التغير في الدخل الشخصي بين الفترة السابقة (t-1) والفترة الحالية (t) هو $100\times\left[\ln(Y_t)-\ln(Y_{t-1})\right]$ وأن البيانات المتوافرة في INCOME.xls هي بيانات لوغاريثمية مسبقاً. ارسم شكلاً بيانياً لبيانات السلسلة الزمنية للسلسلة الجديدة التي أنشأتها للتو.

المدرج التكراري (Histogram):

فيما يتعلق ببيانات السلاسل الزمنية، فإن الرسم البياني الذي يوضح الكيفية التي يتطور بها المتغير على مدى فترة زمنية محددة، غالباً ما يكون أكثر توضيحاً. إلا أنه في حالة البيانات المقطعية، فإن مثل هذه الوسائل (رسم السلاسل الزمنية) لا تكون مناسبة ومن ثم يجب علينا عرض هذه البيانات بطرق أو وسائل أخرى.

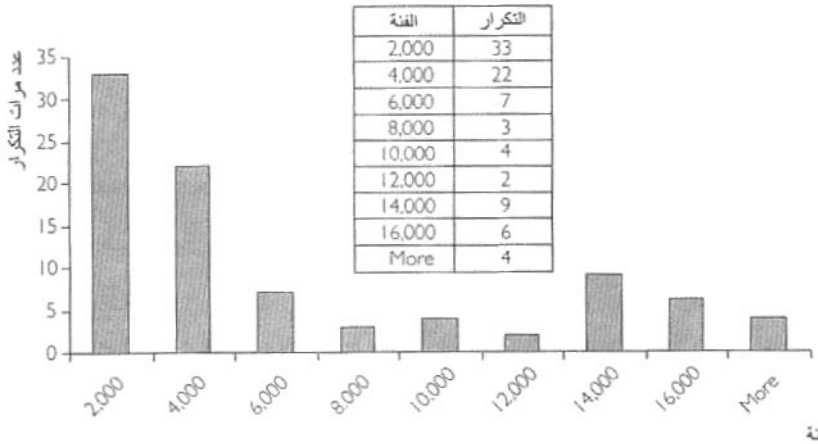
يشتمل ملف إكسل (GDPPC.xls) على بيانات مقطعية عن متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي في عام 1992م لعدد 90 دولة مأخوذة من جداول بيانات (PWT). وقد تم تحويل متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي في كل دولة إلى الدولار الأمريكي باستخدام أسعار الصرف المكافئة للقوة الشرائية بالنسبة لعملة كل دولة من هذه الدول. وهذا يُمكننا من عقد مقارنة مباشرة بين الدول محل الدراسة.

ويعد المدرج التكراري (Histogram) إحدى الوسائل المناسبة لتلخيص تلك البيانات. ولإنشاء المدرج التكراري، فإنه يتوجب عليك أن تبدأ بعمل فئات (Class Intervals) والتي تقسم الدول في شكل مجموعات استناداً إلى متوسط دخل الفرد من إجمالي الناتج المحلي. وفي مجموعة البيانات التي بحوزتنا فإن متوسط دخل الفرد من إجمالي الناتج المحلي يتراوح من 408 دولارات أمريكية في دولة تشاد إلى 17945 دولاراً أمريكياً في الولايات المتحدة الأمريكية. ومن المجموعات التي يمكن الأخذ بها ما يلي: (صفر- 2000)، (2001- 4000)، (4001- 6000)، (6001- 8000)، (8001- 10000)، (10001-12000)، (12001- 14000)، (14001- 16000)، (16001 وأعلى) (وكل الأرقام بالدولار الأمريكي).

لاحظ أن كل فئة (ماعدا فئة 16001 وأعلى) هي بفارق 2000 دولار. بمعنى آخر، أن الفارق في كل فئة من فئاتنا هو 2000. ويمكننا حساب عدد الدول التي تقع في كل فئة بحسب متوسط دخل الفرد من إجمالي الناتج المحلي. على سبيل المثال، هناك سبع دول في فئة بياناتنا تقع ضمن الفئة 4001 دولار أمريكي - 6000 دولار أمريكي. ويسمى عدد الدول الواقعة في الفئة بتكرار (Frequency) تلك الفئة [3]. المدرج التكراري هو رسم بياني على شكل أعمدة تبين التكرارات (Frequencies) لكل الفئات [4].

الشكل رقم (2-2) هو مدرج تكراري لبياناتنا المقطعية عن متوسط دخل الفرد من إجمالي الناتج المحلي التي تستخدم الفئات التي حددت في الفقرة السابقة. لاحظ أنه إذا لم تقم بتحديد الفئة، فإن برنامج إكسل سوف يقوم بتحديدها تلقائياً. برنامج إكسل يقوم أيضاً بإعداد الجدول التكراري (Frequency Table)، الذي يقع فوق المدرج التكراري.

الجدول التكراري يشير إلى عدد الدول التي تتبع لكل فئة. أما الأعداد في العمود المسمى الفئة (Bin) فتشير إلى الحدود العليا لكل فئة.



الفئة	التكرار
2,000	33
4,000	22
6,000	7
8,000	3
10,000	4
12,000	2
14,000	9
16,000	6
More	4

شكل رقم (2-2) مدرج تكراري

على سبيل المثال، يمكننا أن نقرأ أن هناك 33 دولة متوسط دخل الفرد فيها من إجمالي الناتج المحلي أقل من 2000 دولار، و22 دولة متوسط دخل الفرد فيها من إجمالي الناتج المحلي أكثر من 2000 دولار أمريكي إلا أنه أقل من 4000 دولار أمريكي..... وهكذا. المربع الأخير يشير إلى أن هناك 4 دول متوسط دخل الفرد فيها من إجمالي الناتج المحلي أكثر من 16000 دولار أمريكي.

هذه المعلومات نفسها رسمت بطريقة مبسطة في المدرج التكراري. الرسم البياني يتيح الفرصة لإعطاء ملخص نظري سريع لتوزيع بيانات متوسط دخل الفرد عبر الدول.

ومن خلال المدرج التكراري، يمكننا ملاحظة أن العديد من الدول فقيرة جداً، وكذلك ملاحظة "مجموعة" من الدول الغنية جداً (على سبيل المثال، 19 دولة لديها متوسط دخل فرد أكبر من 12000 دولار أمريكي). وهناك دول قليلة نسبياً تقع بين هاتين المجموعتين الفقيرة والغنية (على سبيل المثال، دول قليلة تقع ضمن النطاقات 8000، 10000، 12000). ويشير الاقتصاديون المهتمون بموضوع النمو الاقتصادي غالباً إلى هذا النوع من تركز الدول إلى مجموعات فقيرة وغنية بظاهرة "الذروات المزدوجة". بمعنى آخر، إذا تخيلنا أن المدرج التكراري على شكل جبل، فإننا يمكن أن نرى الذروة عند المجموعة المعنونة 2000 وذروة أصغر عند 14000. طبيعة هذه البيانات يمكن مشاهدتها بسهولة من المدرج التكراري، إلا أنه قد يكون من الصعب إدراك ذلك بسهولة من خلال النظر للبيانات الأولية أو غير المصنفة.

تمرين رقم (2-2):

(أ) أعد رسم المدرج التكراري في الشكل رقم (2-2).

(ب) ارسم مدرجات تكرارية مستخدماً فئات مختلفة. على سبيل المثال، ابدأ بجعل برامج الحاسب الآلي الجاهزة تختار لك قيم الفئات وانظر ماذا ستجد، ثم قم بنفسك باختيار قيم أخرى.

(جـ) إذا كنت تستخدم برنامج إكسل، أعد إجابة السؤالين (أ) و (ب) مع اختيار مربع "Cumulative Percentage" إلى ماذا يؤدي ذلك؟

شكل الانتشار (XY plot):

يهتم الاقتصاديون عادة بشكل العلاقات بين متغيرين أو أكثر، مثال ذلك: هل ترتبط مستويات التعليم العالي والخبرة العملية بالأجور العالية بين العاملين في قطاع معين؟ هل التغيرات في عرض النقود مؤشر يمكن الاعتماد عليه لتفسير التغيرات في مستوى التضخم؟ هل تفسر الاختلافات في حجم الاستثمار لماذا تنمو اقتصاديات بعض الدول بشكل أسرع من غيرها؟

تُعد الأساليب التي ذكرت سابقاً مناسبة لوصف العلاقة الخاصة بسلوك متغير واحد فقط؛ على سبيل المثال، تشير الخصائص المتعلقة بمتوسط دخل الفرد المقطعي من إجمالي الناتج المحلي الحقيقي في الشكل رقم (2-2) إلى أنها غير مناسبة لاختبار العلاقات بين متغيرين.

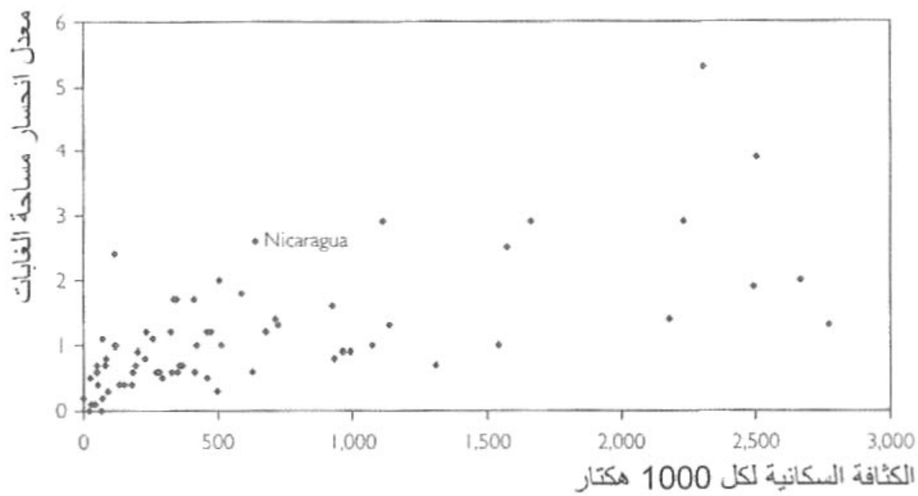
ومادمنا مهتمين بفهم طبيعة العلاقات بين متغيرين أو أكثر من المتغيرات، فإنه من الصعب استخدام الرسم البياني. في الفصول القادمة سوف نناقش تحليل الانحدار (Regression analysis)، الذي يعد الأداة الرئيسة المستخدمة من قبل

الاقتصاديين التطبيقيين الذين يتعاملون مع متغيرات متعددة. ومع ذلك، فإنه يمكن استخدام أساليب الرسوم البيانية لتحديد بعض النواحي البسيطة المتعلقة بالعلاقة بين متغيرين. نقاط انتشار XوY (تسمى أيضاً شكل انتشار- Scatter Diagram) هي مفيدة بشكل خاص في هذا الجانب.

الشكل رقم (2-3) هو عبارة عن رسم بياني لبيانات حول معدل انحسار مساحة الغابات (Deforestation) (المتوسط السنوي للمساحة المفقودة من الغابات خلال الفترة 1981 ــ 1990م كنسبة مئوية من إجمالي المساحة التي تشغلها الغابات لعدد 70 دولة استوائية)، مع بيانات كثافة السكان بها (أي، عدد الأشخاص في كل ألف هكتار Population per 1000 hectars). (هذه البيانات متوافرة في الملف FOREST.xls) إذ هناك اعتقاد كبير أن الدول ذات الكثافة السكانية العالية تكون أكثر قابلية لانحسار مساحة الغابات بشكل أكثر من الدول ذات الكثافة السكانية المنخفضة، باعتبار أن الكثافة السكانية العالية قد تؤدي إلى زيادة الضغط لقطع الأشجار لاستخدامها خشباً للوقود أو لتخصيص مساحات الغابات لتكون أراض زراعية مطلوبة لزراعة المحاصيل الزراعية.

الشكل رقم (2-3) هو شكل نقاط انتشار X Y لهذين المتغيرين. كل نقطة على هذا الشكل تمثل دولة معينة. قراءتنا لمحور Y (المحور الرأسي) توضح لنا معدل انحسار مساحة الغابات في تلك الدولة. أما قراءتنا لمحور X (المحور الأفقي) فإنها توضح لنا الكثافة السكانية لكل 1000 هكتار. حيث يكون من الممكن تسمية كل نقطة باسم الدولة التي تنطبق عليها البيانات التي تشير إليها تلك النقطة. ولم نقم بذلك في الشكل رقم (2-3)؛ لأن تسمية 70 دولة سوف يجعل الشكل مزدحماً وتصعب قراءته. وبالرغم من ذلك، فقد قمنا بتسمية النقطة الخاصة بدولة نيكارجوا. لاحظ أن هذه الدولة لديها معدل انحسار لمساحة

الغابات يصل إلى 2.6%سنوياً (Y=2.6) وكثافة سكانية تصل إلى 640 شخصاً لكل ألف هكتار (X=640).



شكل رقم (3-2): نقاط التقاء YوX للكثافة السكانية مقابل معدل الانحسار في مساحة الغابات

ويمكن استخدام شكل الانتشار لإعطاء لمحة سريعة عن العلاقة بين معدل انحسار مساحة الغابات والكثافة السكانية. فتحليل هذا الشكل يشير إلى تأكيد فكرة العلاقة بين معدل انحسار مساحة الغابات والكثافة السكانية. فمثلاً، إذا نظرنا للدول ذات الكثافة السكانية المنخفضة، (أقل من 500 شخص لكل هتكار) فسنجد معظمها ذات معدلات منخفضة في انحسار مساحة الغابات (أقل من 1% سنوياً). وإذا نظرنا للدول ذات الكثافات السكانية العالية (على سبيل المثال، أكثر من 1500 شخص لكل ألف هكتار) فسنجد معظمها ذات معدلات عالية في انحسار مساحة الغابات (أكثر من 2% سنوياً) مما يشير إلى أنه قد يكون هناك علاقة إيجابية بين الكثافة السكانية ومعدل انحسار مساحة الغابات (بمعنى آخر، أن القيم العالية لأحد المتغيرين تميل للارتباط بالقيم العالية للمتغير الآخر، والقيم المنخفضة ترتبط بالقيم المنخفضة). كما أنه من الممكن أيضاً أن نجد علاقة سالبة

بين المتغيرين. فقد يحدث ذلك، إذا استخدمنا معدل التحضر (Jrbanization) بديلاً عن الكثافة السكانية في شكل الانتشار. ففي هذه الحالة، فإن مستويات التحضر العالية قد تكون مرتبطة بمعدلات منخفضة لانحسار مساحة الغابات مادام أن التوسع في المدن قد يخفض الضغوطات السكانية في المناطق الريفية حيث تتواجد الغابات.

يجدر بالذكر أن العلاقات الموجبة أو السالبة التي توجد في البيانات ما هي إلا مؤشرات أو نزعات، وعلى ذلك فإنها ليست بالضرورة أن تنطبق على كل دولة. أي أنه قد يكون هناك استثناءات للنمط العام للكثافة السكانية العالية التي ترتبط بمعدلات عالية لانحسار مساحة الغابات. على سبيل المثال، في شكل الانتشار يمكننا ملاحظة دولة واحدة ذات كثافة سكانية عالية تصل إلى 1300 تقريباً ولديها في نفس الوقت معدل منخفض في انحسار مساحة الغابات يصل إلى 0.7%. وفي المقابل، فإن الكثافة السكانية المنخفضة يمكن أن يصاحبها أيضاً معدلات عالية في انحسار مساحة الغابات، بحسب ما تم إثباته في دولة ذات كثافة سكانية منخفضة تصل إلى 150 تقريباً إلا أنها ذات معدل عال في انحسار مساحة الغابات يصل إلى 2.5 % سنوياً تقريباً. وبوصفنا اقتصاديين، فإننا عادة ما نكون مهتمين بتحديد الأنماط العامة للبيانات. ومع ذلك، فإنه يجب علينا أن نضع في الاعتبار دائماً وجود تلك الاستثناءات (وهي تسمى قيم شاذة أو قيم خارجة عن السياق في لغة الإحصائيين) لهذه الأنماط. وفي بعض الحالات فإن معرفة الدول التي لا تتفق بياناتها مع النمط العام تكون أيضاً مثيرة للاهتمام والدراسة.

تمرين رقم (2-3):

يشتمل الملف FOREST.xls على بيانات حول كل من نسبة الزيادة في مساحة أراضي زراعة المحاصيل (العمود المعنون "Crop ch") من 1980 إلى 1990م وعن نسب الزيادة في المراعى الدائمة (العمود المعنون "Pasture ch") خلال نفس الفترة. ارسم شكل انتشار لهذين المتغيرين (كل متغير على حدة) في مقابل معدل انحسار مساحة الغابات. هل يبدو أن هناك علاقة إيجابية بين معدل انحسار مساحة الغابات والتوسع في الأراضي المخصصة للمراعي؟ وما هو شكل العلاقة بين معدل انحسار مساحة الغابات والتوسع في الأراضي المخصصة لزراعة المحاصيل؟.

التعامل مع البيانات: الإحصاءات الوصفية:

تحتوي الرسوم البيانية على تأثير تصويري مباشر يساعد في إثراء مقالة أو تقرير حول الموضوع مدار البحث. إلا أن العديد من الحالات تتطلب دقة في التحليل الرقمي للموضوع محل البحث. وستقوم الفصول التالية بوصف الأساليب الرقمية شائعة الاستخدام لتلخيص العلاقة بين العديد من المتغيرات بشكل أكثر تفصيلاً. وهنا سوف نناقش باختصار بعض الإحصاءات الوصفية التي تصف متغير واحد. ومن باب التحفيز، سوف نرجع لمفهوم التوزيع الذي ذكرناه في نقاشنا حول المدرجات التكرارية.

في بياناتنا المقطعية عن متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي لعدد 90 دولة، نلاحظ أن التغير فيها يتضح من خلال النظر إلى المدرج التكراري في الشكل رقم (2-2) الذي يوضح توزيع متوسط دخل الفرد من إجمالي الناتج المحلى في هذه الدول. افترض أنك تريد تلخيص المعلومات الموجودة في المدرج

التكراري بشكل رقمي، فما عليك سوى أخذ الأرقام المعروضة في الجدول في الشكل رقم (2-2). ومع ذلك، حتى هذا الجدول قد يقدم أرقاماً عديدة كثيرة لا يمكن تفسيرها بسهولة. وبدلا عن ذلك نستخدم رقمين بسيطين هما المتوسط الحسابي (mean) والانحراف المعياري (standard deviation).

المتوسط الحسابي هو مصطلح إحصائي للمتوسط، والصيغة الرياضية للمتوسط الحسابي هى:

$$\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}$$

حيث N هي حجم العينة (بمعنى آخر عدد الدول) والعلامة Σ هي علامة الجمع (بمعنى آخر، أنها تجمع متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي لكل الدول). وفي هذا المثال فإن متوسط دخل الفرد من إجمالي الناتج المحلي هو 5.443.80 دولاراً أمريكياً.

في جميع فصول هذا الكتاب، سوف نقوم بوضع شرطة (bar) على المتغير للإشارة إلى متوسطه الحسابي (بمعنى آخر، $\bar{Y}$ هي المتوسط الحسابي للمتغير Y، $\bar{X}$ هي المتوسط الحسابي للمتغير X وهكذا).

مفهوم المتوسط الحسابي يرتبط بمتوسط التوزيع. على سبيل المثال إذا نظرنا للمدرج التكراري السابق نجد أن الرقم 5.443.80 دولاراً أمريكياً يقع في متوسط التوزيع.

توزيع البيانات المقطعية لمتوسط دخل الفرد من إجمالي الناتج المحلي هو توزيع غير عادي، لما لديه من خاصية الذروة المزدوجة التي تم وصفها سابقاً.

الأكثر شيوعاً أن يكون لتوزيعات المتغيرات الاقتصادية ذروة واحدة وأن تأخذ شكل الجرس (bell-shaped).

الشكل رقم (2-4) هو عبارة عن مدرج تكراري يوضح التوزيع الذي يأخذ شكل الجرس. ولمثل هذه التوزيعات يكون موقع المتوسط الحسابي بالتحديد في وسط التوزيع أسفل ذروة البيانات.

وبالطبع، فإن المتوسط الحسابي يخفي قدراً كبيراً من الاختلاف والتغاير بين الدول. وتعتبر القيم الدنيا والعليا من الملخصات الإحصائية المفيدة التي تسلط الضوء بشكل أكبر على الاختلافات بين الدول في معدلات دخل الفرد من إجمالي الناتج المحلي. وبالنسبة لبياناتنا، فإن أدنى متوسط دخل فرد هو 408 دولارات أمريكية لدولة تشاد، وأعلى متوسط دخل فرد هو 17945 دولاراً أمريكياً في الولايات المتحدة الأمريكية. ومن خلال النظر إلى الفارق بين الأعلى والأدنى يمكننا معرفة تشتت توزيع البيانات.

يُعد مفهوم التشتت مهماً في علم الاقتصاد ويتعلق بشكل مباشر بمفاهيم الاختلاف وعدم المساواة. على سبيل المثال، يتراوح متوسط دخل الفرد في بياناتنا من 408 إلى 17945 دولارا أمريكياً.

وإذا نمت الدول الأفقر بشكل سريع في المستقبل القريب، وبقيت الدول الأغنى على حالها، فإن توزيع متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي في عام 2012م مثلاً قد يقلص فجوة التشتت في الدخل بين مجموعتي الدول هاتين. وقد يكون الوضع هو أن يصل متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي إلى 10000 دولار أمريكي في أفقر دولة، في حين يبقى ذلك المتوسط في الدول الغنية عند مستوى 17945 دولاراً أمريكياً. فإذا كان ذلك ما سوف يحدث فإن توزيع بيانات متوسط دخل الفرد من إجمالي الناتج المحلي الحقيقي يكون أكثر عدالة (أقل تشتتاً، أقل اختلافاً).

ومن البديهي أن يوجد ترابط بين مفاهيم التشتت، الاختلاف/ التغاير، وعدم العدالة/ المساواة.

لكن لايمكننا الاعتماد بشكل كبير على القيم الدنيا والقيم العليا كمؤشر موثوق لدرجات التشتت. على سبيل المثال، ماذا يحدث ــ باستثناء دولة تشاد إذا شهدت كل الدول الفقيرة نمواً اقتصادياً سريعاً بين عامي، 1992م و 2012م، في حين لم يحدث أي نمو اقتصادي في الدول الغنية على الإطلاق؟ ففي هذه الحالة سوف تنخفض درجات التشتت وعدم المساواة بين الدول بمرور الزمن. وعلى ذلك، فإذا لم تنمو كل من دولة تشاد والولايات المتحدة الأمريكية فإن القيم الدنيا والقيم العليا سوف تبقى على حالها عند 408 دولارات أمريكية و 17945 دولاراً أمريكياً على التوالي.

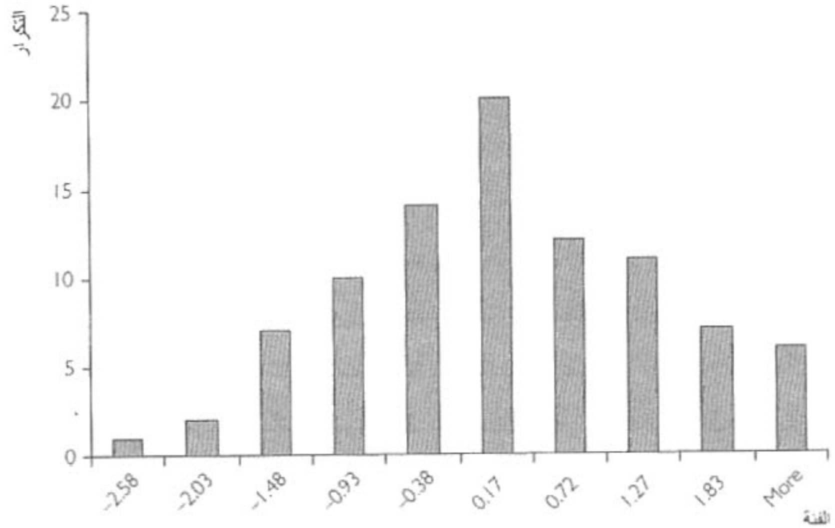
يعد الانحراف المعياري (standard deviation) من أهم مقاييس التشتت المستخدمة (وبطريقة أكثر أرباكاً، يسمي الإحصائيون مربع قيمة الانحراف المعياري بالتباين variance)

ويتم التعبير عن الانحراف المعياري بالصيغة الرياضية التالية:.

$$s = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^2}{N-1}}$$

ومن الملاحظ أنك عملياً لا تحتاج لحسابه يدوياً، حيث يمكنك حسابه بسهولة في برنامج إكسل باستخدام خاصية الدوال الرياضية المبرمجة مسبقاً.

وفي بعض الكتب الأساسية، توجد صيغة رياضية مختلفة قليلاً لحساب الانحراف المعياري حيث يتم وضع N في المقام بدلاً عن N-1.



شكل رقم (2-4) المدرج التكراري

هذا المقياس (الانحراف المعياري) قد يصعب استيعابه للوهلة الأولى في بياناتنا عن إجمالي الناتج المحلي؛ فقيمة الانحراف المعياري هي 5369.496 دولاراً أمريكياً قد يصعب وضع تفسير واضح لما يعنيه هذا الرقم بشكل مطلق. ومع ذلك، فإنه يمكن التعبير عن الانحراف المعياري من خلال المقارنة، أي يمكنك مقارنة الانحرافات المعيارية لتوزيعين مختلفين، حيث غالباً ما يعرض التوزيع الأقل انحرافاً تشتتاً أقل. في مثالنا أعلاه، إذا شهدت الدول الفقيرة نمواً اقتصادياً مفاجئاً، وبقي الحال على ما هو عليه في الدول الغنية، فإن الانحراف المعياري سوف ينخفض مع مرور الوقت.

تمرين رقم (2-4):

قم باحتساب وتفسير الإحصاءات الوصفية للتغير في المتغيرين: مساحة المراعي ومساحة الأراضي المخصصة لزراعة المحاصيل في الملف FOREST. XLS

ملخص الفصل:

1- البيانات الاقتصادية تكون في عدة أشكال. وتعد بيانات السلسلة الزمنية، والبيانات المقطعية وبيانات السلاسل الزمنية المقطعية من أكثر أنواع البيانات شيوعاً.

2- يمكن الحصول على البيانات الاقتصادية من عدة مصادر. وقد أصبحت شبكة الإنترنت مستودعاً ذا قيمة متزايدة لكثير من البيانات.

3- تعتبر أساليب الرسم البياني البسيطة بما فيها المدرجات التكرارية وأشكال الانتشار من الطرق المفيدة لتلخيص المعلومات التي تحتويها البيانات.

4- هناك العديد من الإحصاءات الوصفية التي يمكن استخدامها. ومن أهم هذه الإحصاءات المتوسط الحسابي باعتباره مقياساً لموقع توزيع البيانات، والانحراف المعياري مقياساً لمدى تشتت هذا التوزيع.

ملحق رقم (2-1): الأرقام القياسية:

لتوضيح أسس إنشاء واحتساب الرقم القياسي للأسعار، فإننا نستخدم البيانات الموضحة في الجدول رقم (2-1-1)عن أسعار فواكه مختلفة ولعدة سنوات.

حساب الرقم القياسي لسعر فاكهة الموز:

نبدأ بحساب الرقم القياسي لسعر فاكهة الموز، ومن ثم نقوم باحتساب رقم قياسي لأسعار الفواكهه ككل. ووفقا لما تم إيضاحه سابقاً، فإن حساب الرقم القياسي يبدأ باختيار سنة الأساس والتي في مثال الموز هي سنة 2000م (مع

التأكيد بأنه يمكن اختيار أي سنة أخرى). وعلى ضوء هذا الاختيار، فإن قيمة الرقم القياسي لفاكهة الموز هو 100 في عام 2000م سنة الأساس. والسؤال الذي يطرح نفسه هو: كيف قمنا بتحويل سعر الموز في عام 2000م للحصول على قيمة الرقم القياسي 100؟ والإجابة هى أن هذا التحويل تم بأخذ سعر الموز فى عام 2000م وقسمته على سعر الموز في العام نفسه (بمعنى آخر قسمة السعر على نفسه) والضرب في 100. ومن أجل المحافظة على إمكانية المقارنة، يجب تطبيق هذا التحويل على سعر الموز في كل سنة للحصول على الرقم القياسي للموز، وقد تم توضيح ذلك في الجدول رقم (2-1-2).

ومن الرقم القياسي للموز، يمكن ملاحظة أن سعر الموز قد ارتفع بنسبة 4.4% خلال الفترة من 2000 إلى 2003م، وفي عام 1999م فإن سعر الموز كان يوازي 97.8% مقارنة بسعره في عام 2000م.

جدول رقم (2-1-1) أسعار فواكه مختلفة في سنوات مختلفة (جنيه/الكيلو)

السنة	الموز	التفاح	فاكهة الكيوى
1999	0.89	0.44	1.58
2000	0.91	0.43	1.66
2001	0.91	0.46	1.90
2002	0.94	0.50	2.10
2003	0.95	0.51	2.25

جدول رقم (2-1-2) حساب الرقم القياسي لسعر الموز

السنة	سعر الموز	التحويل	الرقم القياسي لسعر الموز
1999	0.89	X100 / 0.91	97.8
2000	0.91	X100 / 0.91	100
2001	0.91	X100 / 0.91	100
2002	0.94	X100 / 0.91	103.3
2003	0.95	X100 / 0.91	104.4

حساب الرقم القياسي لأسعار الفاكهة:

عند حساب الرقم القياسي لسعر الموز (سلعة واحدة)، فكل ما يجب علينا عمله هو النظر إلى أسعار الموز في السنوات المختلفة. أما إذا أردنا حساب الرقم القياسي لأسعار مجموعة من الفواكه (سلع مختلفة) فإنه يجب علينا إيجاد طريقة لتوحيد الأسعار لكل أنواع الفاكهة مع بعضها. أحد الطرق التي يمكن اتباعها هو أخذ المتوسط لأسعار كل الفواكه مع بعضها في كل سنة (وبعد ذلك يتم حساب رقم قياسي بنفس الطريقة التي تم بها حساب الرقم القياسي للموز). ولكن هذه الطريقة عادة ما تكون غير مناسبة لأنها تساوي ضمنياً كل السلع مع بعضها (بمعنى آخر، أن المتوسط البسيط هو عبارة عن جمع أسعار الفواكه الثلاثة وقسمتها على ثلاثة). في مثالنا (وفي معظم التطبيقات العملية)، فإن هذه المساواة بين الأسعار غير مناسبة. (الاستثناء لذلك هو متوسط داوجونز الصناعي DJIA الذي يوازن أو يساوي بين أوزان أسعار الأسهم لكل الشركات التي يتكون منها المؤشر).

يوضح جدول رقم (2-1-1) أن أسعار الموز والتفاح تتصاعد بشكل طفيف بمرور الزمن (وأنه في بعض السنوات لاتتغير أسعارهما وربما تنخفض)، في حين أن سعر فاكهة الكيوى يتزايد بشكل ملحوظ بمرور الوقت. الموز والتفاح فاكهتان يشتريهما غالبية الأفراد باستمرار، في حين أن فاكهة الكيوى غريبة وغير معروفة لعامة الأفراد ويتم شراؤها بصورة متقطعة من قبل فئة قليلة من الأفراد. وفى ضوء ذلك، فإنه من غير الملائم مساواة أوزان أسعار الفواكه الثلاثة عند حساب الرقم القياسي لأسعار الفواكه. الرقم القياسي لأسعار الفاكهة المبني على المتوسط البسيط سوف يكشف بأن أسعار الفواكه كانت تنمو بمعدل معتدل السرعة (بمعنى آخر، أن الجمع بين النمو المنخفض لسعري الموز والتفاح مع النمو السريع جدا لسعر فاكهة الكيوي سوف تكون نتيجته سعر فاكهة قياسي يشير إلى نمو في السعر متوسط السرعة)، وعلى الرغم من ذلك، إذا كان يتوجب على الحكومة استخدام مثل هذا الرقم القياسي للإشارة إلى أن "أسعار الفواكه تتزايد بمعدل سريع" فإن معظم الأفراد سيجدون أن ذلك لايتماشي مع تجربتهم الشخصية. أى أن معظم الأفراد الذين يشترون الموز والتفاح فقط، سوف يجدون أن الأسعار لا تتغير بشكل سريع وملحوظ.

وتبعاً للمنطق، الذي ورد في الفقرة السابقة، فإن الرقم القياسي الذي يوازن بين كل السلع بطريقة متساوية لن يكون ذا معنى أو مقبولاً. كما يشير إلى الكيفية التى يمكن بواسطتها حساب رقم قياسي مقبول لأسعار الفاكهة كالتالي: استخدم متوسط مرجح لأسعار جميع الفواكه بحيث يتم اختيار الأوزان لتعكس أهمية كل سلعة. وفي مثالنا عن الرقم القياسي للفاكهة، فإننا نريد أن نضع وزناً أكثر لفاكهة الموز و التفاح (أكثر الفواكه شيوعاً) ووزناً أقل لفاكهة الكيوي (الغريبة) (5). وهناك العديد من الطرق لاختيار مثل هذه الأوزان وسوف يتم هنا شرح أسلوبين شائعين يقومان على فكرة أن الأوزان يجب أن تعكس الكمية التى يتم شراؤها من

كل فاكهة. وبالطبع، فإن الكمية المشتراة من كل فاكهة، قد تختلف عبر الزمن، وبأخذ ذلك في الحسبان، فإن السعرين القياسين اللذين قمنا بحسابهما يختلفان.

الرقم القياسي للأسعار لاسبير (باستخدام أوزان سنة الأساس):

يستخدم الرقم القياسي للأسعار لاسبير ((The Lasperes Price index (LPI) كمية كل فاكهة تم شراؤها في سنة الأساس (عام 2000م في مثالنا) لحساب الأوزان. ومعنى ذلك أنه لكي نتمكن من حساب الرقم القياسي للأسعار لاسبير، فإنه يتوجب عليك حساب متوسط سعر الفاكهة في كل سنة مستخدماً المتوسط المرجح، إذ إن الأوزان تتناسب مع كمية كل فاكهة تم شراؤها في عام 2000م. بعد ذلك استخدم متوسط سعر الفاكهة هذا لحساب رقم قياسي بنفس الطريقة التي قمنا بها عند حساب الرقم القياسي للموز (انظر جدول رقم 2-1-2).

ومن البديهى، إذا أنفق المستهلك العادى على فاكهة الموز 100 مرة أكثر مما أنفقه على فاكهة الكيوى في عام 2000م، فإن أسعار الموز ستحصل على وزن يفوق 100 مرة الوزن الذي ستحصل عليه أسعار الكيوي في الرقم القياسي لاسبير. ويمكن التعبير عن الرقم القياسي لاسبير كما يلي: اجعل P تمثل سعر السلعة، Q تمثل الكمية المشتراة من السلعة، والأرقام السفلية تمثل السلعة والسنة، فالموز هو السلعة 1، التفاح السلعة 2، الكيوي هو الفاكهة 3. وبذلك، على سبيل المثال فإن P1,2000 هو سعر الموز في عام 2000م، Q3,2002 هي الكمية المشتراة من فاكهة الكيوى في عام 2002م ... الخ، وإذا وجدت صعوبة في فهم هذا الترقيم أو معامل الجمع المستخدم أدناه، انظر الملحق رقم (1-1).

ووفقا للنقاش السابق، فإن الرقم القياسي للأسعار لاسبير (LPI) في السنة t، (t=1999,2000, 2001, 2002 and 2003) يمكن كتابته على النحو التالي:

$$LPI_t = \frac{\sum_{i=1}^{3} P_{it} Q_{i,2000}}{\sum_{i=1}^{3} P_{i,2000} Q_{i,2000}} x100$$

لاحظ أن البسط في هذه المعادلة يأخذ سعر كل فاكهة ويضربها في الكمية المشتراة من تلك الفاكهة في العام 2000م. وهذا يضمن أن الموز والتفاح يكون لهما وزن أكبر في الرقم القياسي للأسعار لاسبير. ولن نقوم بتوضيح المقام أكثر من القول أنه من الضروري التأكيد على أن الرقم القياسي للأسعار لاسبير هو رقم قياسي صحيح مع وجود قيمة الـ 100 في سنة الأساس. أما بالنسبة للقارئ الذي يميل للتفسير الرياضي، فإن المقام يضمن لنا أن الأوزان الموجودة في المتوسط المرجح مجموعها واحد (وهو أمر ضروري للتأكيد بأنه المتوسط المرجح المناسب).

لاحظ أيضاً أن التعريف الوارد أعلاه للرقم القياسي للأسعار لاسبير قد تمت كتابته أو صياغته لمثالنا عن الفواكه والذي يتضمن ثلاث سلع بسنة أساس هي عام 2000م. بشكل عام، فإن المعادلة أعلاه يمكن توسيعها للسماح لأي عدد من السلع ولأي سنة أساس من خلال تغيير الرقم3 والرقم 2000 بحسب الحالة.

طريقة حساب الرقم القياسي للأسعار لاسبير تتطلب منا أن نعرف الكميات المشتراة من كل فاكهة، والجدول رقم (2-1-3) يوضح هذه الكميات.

ويمكن تفسير الرقم القياسي لأسعار الفواكه لاسبير بنفس طريقة تفسير الرقم القياسي لسعر الموز. على سبيل المثال، يمكننا القول بأن أسعار الفواكه ارتفعت بين عامي 2000م و2003م بنسبة 8.7%.

الرقم القياسي للأسعار لاباستشي (باستخدام أوزان السنة الحالية):

استخدم الرقم القياسي للأسعار لاسبير أوزان سنة الأساس لحساب متوسط أسعار الفواكه من أسعار الأنواع الثلاثة من الفواكه. إلا أنه من الممكن أن تكون أوزان سنة الأساس (في مثالنا سنة الأساس كانت عام 2000م) غير مناسبة إذا كانت أنماط استهلاك الفواكه تتغير بشكل ملحوظ مع مرور الوقت. في مثالنا يلاحظ أن الموز والتفاح هما الفاكهتان الأكثر مبيعاً في كل السنوات، وأن هناك القليل من مستهلكي فاكهة الكيوى.

جدول رقم (2-1-3) الكميات المشتراة من الفواكه (آلاف الكيلو جرامات)

السنة	الموز	التفاح	فاكهة الكيوي
1999	100	78	1
2000	100	82	1
2001	98	86	3
2002	94	87	4
2003	96	88	5

جدول رقم (2-1-4) حساب الرقم القياسي لأسعار الفواكه (لاسبير)

السنة	البسط $\sum_{i=1}^{3} P_{it} Q_{i2000}$	المقام $\sum_{i=1}^{3} P_{i,2000} Q_{i,2000}$	الرقم القياسي للأسعار (LPI)
1999	126.64	127.92	99.00
2000	127.92	127.92	100.0
2001	130.62	127.92	102.1
2002	137.1	127.92	107.2
2003	139.07	127.92	108.7

الرقم القياسي للأسعار حسب طريقة لاسبير يعطي وزناً أكبر لأسعار الموز والتفاح بطريقة (ذات معنى) مقارنة بأسعار فاكهة الكيوى، إلا أنه ماذا كان سيحدث في عام 2001م، إذا كان هناك تخوف صحي من أن أكل التفاح يضر بالصحة وأن الأفراد امتنعوا عن أكل التفاح وبدلاً عن ذلك أكلوا الكثير من فاكهة الكيوى؟ الإجابة هي أن الرقم القياسي للأسعار لاسبير سوف يواصل تخصيص وزناً منخفضاً لفاكهة الكيوى ووزناً أعلى للتفاح حتى مع وجود حقيقة أن الأفراد أصبحوا يأكلون الآن فاكهة الكيوى أكثر. لذلك، يحاول الرقم القياسي للأسعار لباستشي ((Paasche Price Index(PPI) تجاوز هذه المشكلة من خلال استخدام مشتريات السنة الحالية لإعطاء وزن لكل فاكهة فيه.

بمعنى آخر، لكي تنشئ الرقم القياسي للأسعار لباستشي فإنه يجب عليك حساب السعر المتوسط للفاكهة في كل سنة باستخدام المتوسط المرجح، حيث إن الأوزان تتناسب مع كمية كل فاكهة يتم شراؤها خلال السنة الحالية. حيث يمكنك بعد ذلك استخدام متوسط سعر الفاكهة هذا لحساب رقم قياسي بنفس الطريقة التي حسبنا بها الرقم القياسي للموز (انظر جدول رقم 2-1-2)

يمكن كتابة المعادلة الرياضية للرقم القياسي للأسعار لباستشى (PPI) في السنة t، (t=1999,.......,2003) على النحو التالي:

$$PPI_t = \frac{\sum_{i=1}^{3} P_{it} Q_{it}}{\sum_{i=1}^{3} P_{i,2000} Q_{it}} x100$$

لا حظ أن (PPI) هو مثل (LPI) ما عدا أن Q_{it} تظهر في معادلة PPI بينما ظهرت Q_{i2000} في معادلة LPI . وبذلك، فإن السعرين القياسيين متساويان عدا

فيما يتعلق بحقيقة أن PPI يستخدم كميات السنة الحالية بدلاً عن كميات سنة الأساس.

يوضح جدول رقم (2-1-5) كيفية حساب الرقم القياسي لباستشي باستخدام بيانات أسعار الفواكه الواردة في الجدول رقم (2-1-1) وبيانات كمية الفواكه المشتراة الواردة في الجدول رقم (2-1-3).

لاحظ أنه ما دام الرقم القياسي للأسعار لباستشي لا يرجح الأسعار بنفس طريقة الرقم القياسي للأسعار لاسبير، فإننا لا نحصل على نفس النتائج الواردة في الجداول (2-1-5 و 2-1-4). على سبيل المثال، يشير الرقم القياسي للأسعار لباستشي إلى أن أسعار الفواكه ارتفعت بنسبة 10.4% بين عامي 2000م و2003م (في حين يشير الرقم القياسي للأسعار لاسبير إلى أن أسعار الفواكه ارتفعت بنسبة 8.7%).

باستشي ولاسبير هما رقمان قياسيان للأسعار من بين عشرات الأرقام القياسية، ولكننا لن نقوم بمناقشة أرقام قياسية أخرى. ومن المهم أن نعلم أن الأرقام القياسية للأسعار يتم استخدامها بكثرة في مجالات علم الاقتصاد والتمويل. على سبيل المثال، تستند قياسات معدل التضخم الذي تنشره الصحف على الأرقام القياسية للأسعار. في الاقتصاد، فإن آلاف السلع التي يمكن أن يشتريها الأفراد، والأسعار القياسية للمستهلك (CPI) أو الأسعار القياسية للتجزئة (RPI) هي عبارة عن متوسطات مرجحة لأسعار هذه الآلاف من السلع. أما المعلومات عن أسواق رأس المال (الأسهم) فغالباً ما يتم التعبير عنها في شكل أرقام قياسية للأسهم.

جدول رقم (2-1-5) الرقم القياسي لأسعار الفواكه حسب طريقة باستشي

السنة	البسط $\sum_{i=1}^{3} P_{it} Q_{it}$	المقام $\sum_{i=1}^{3} P_{2000} Q_{it}$	الرقم القياسي لباستشي (PPI)
1999	124.90	126.20	99.0
2000	127.92	127.92	100.0
2001	134.44	131.14	102.5
2002	140.26	129.59	108.2
2003	147.33	133.50	110.4

جدول رقم (2-1-6) الربط بين الأرقام القياسية للأسعار عندما تختلف سنة الأساس

السنـة	الرقم القياسي القديم لسنة الأساس **1995**	الرقم القياسي الجديد لسنة الأساس **2001**	التحويل للرقم القياسي القديم	الرقم القياسي لسنة الأساس **2001**
1995	100			88.8
1996	102			90.6
1997	103		X95 / 107	91.5
1998	103		X95 / 107	91.5
1999	105		X95 / 107	93.2
2000	107	95	X95 / 107	95
2001		100	X95 / 107	100
2002		101		101
2003		105		105

هناك قضية أخرى تعرقل في بعض الأحيان الدراسات التطبيقية، وبخاصة تلك الدراسات التى تتضمن بيانات عن الاقتصاد الكلي. غالباً ماتقوم الأجهزة الحكومية الإحصائية بتحديث سنة الأساس التى تستخدمها في حساب الأرقام القياسية للأسعار. لذلك، فعندما تقوم بجمع البيانات قد تواجه بحالة يكون فيها الجزء الأول من بياناتك يستخدم سنة أساس معينة والجزء الأخير من البيانات يستخدم سنة أساس مختلفة. هذه المشكلة يسهل حلها إذا كان لديك سنة تتداخل فيها بيانات الرقمين (Overlap) بحيث تعرف الرقم القياسي للأسعار على أساس كل من سنتي الأساس. الجدول رقم (2-1-6) يقدم توضيحاً للكيفية التى يمكن لك من خلالها وصل (splicing) الأرقام القياسية مع بعضها عندما تتغير سنة الأساس بهذه الطريقة.

قام المكتب الإحصائي بحساب أرقام قياسية مستخدما سنة 1995م كسنة أساس، ولكن توقف عن إصدار هذا الرقم بعد عام 2000م. وهذا يمكن قراءته في العمود المعنون: "الرقم القياسي القديم لسنة الأساس 1995".

في عام 2001م، بدأ المكتب الإحصائي في حساب الأرقام القياسية للأسعار مستخدماً عام 2001م كسنة أساس، إلا أنه أعاد حساب الرقم القياسي الخاص بسنة 2000م لهذه الأسعار مستخدماً سنة الأساس الجديدة.

هذا الرقم الجديد تم وضعه في العمود المعنون: "الرقم القياسي الجديد لسنة الأساس 2001".

لاحظ أنه لدينا سنة واحدة تتداخل فيها البيانات، وهى سنة 2000م. ومن أجل أن نجعل الرقم القياسي لسنة 2000 القديم يساوي الرقم القياسي الجديد يجب علينا أن نأخذ قيمة الرقم القياسي القديم ونضربها في 95 ثم نقسمها على 107. ومن أجل أن يكون عملنا في سياق واحد، فإننا يجب أن نطبق نفس هذا التحويل لكل القيم الواردة في الرقم القياسي القديم. ناتج تحويل كل قيم الرقم القياسي القديم

بهذه الطريقة موجود في العمود الأخير من الجدول رقم (2-1-6). هذا الرقم القياسي يمكن الآن استخدامه في عمل بحثي تطبيقي؛ لأن الرقم القياسي كله الآن أصبح له نفس سنة الأساس وهى سنة 2001م.

ملحق رقم (2-2): الإحصاءات الوصفية المتقدمة:

على الرغم من كثرة الإحصاءات وتعددها إلا أن المتوسط الحسابي والانحراف المعياري هما من أكثر الإحصاءات الوصفية شيوعاً، فالمتوسط الحسابي هو أبسط مقياس لتحديد موقع التوزيع. كلمة "موقع" (Location) استخدمت لتعكس فكرة مركز التوزيع. المقاييس الأخرى الشائعة لتحديد "الموقع" هى المنوال (Mode) والوسيط (Median).

وللتفريق بين المتوسط الحسابي، والمنوال والوسيط اقرأ المثال المبسط التالي. سبعة أشخاص أفادوا بأن دخلهم السنوي بالجنيه الأسترلينى هو: 18000، 15000، 9000، 15000، 16000، 17000، 20000. المتوسط الحسابي، أو المتوسط، لدخل هؤلاء الأشخاص السبعة هو 15.714 جنيهاً إسترلينياً. الوسيط هو القيمة التى تتوسط التوزيع. أى أنها القيمة التى تقسم التوزيع إلى نصفين متساويين. وفى مثالنا أعلاه، هو قيمة الدخل التى عندها نجد نصف هؤلاء الأشخاص دخولهم عالية ونصفهم دخولهم منخفضة. والوسيط هنا هو 16000 جنيه إسترليني. لاحظ أن ثلاثة أشخاص دخولهم أقل من قيمة الوسيط وأن ثلاثة أشخاص دخولهم أعلى من قيمة الوسيط.

يمكن أيضا فهم المزيد حول خصائص كل من المنوال والوسيط من خلال ماورد في الشكلين (2-2 و 2-4) اللذين يظهران مدرجين تكراريين أو توزيعين. مشكلة المنوال تتمثل في أنه قد لايكون هناك قيمة واحدة دارجة بين البيانات. على سبيل المثال، في مجموعة بيانات متوسط دخل الفرد من إجمالي الناتج

المحلي (GDPPC.XLS)، ليس هناك دولتان لديهما نفس مستوى الدخل؛ لذلك ليس هناك قيمة تكررت أكثر من مرة.

وعادة يكون المنوال في مثل هذه الحالات أعلى نقطة في المدرج التكراري. هناك مشكلة بسيطة تتعلق بتعريف المنوال بهذه الطريقة تتمثل في أنه قد يكون حساساً لمسألة اختيار فترات المجموعة (وهذا يبرر لماذا يعطي برنامج إكسل إجابة تختلف قليلاً بالنسبة للمنوال لبيانات GDPPC.XLS مقارنة بما هو معطى هنا). في الشكل رقم (2-2) المدرج التكراري يكون أعلى من مجموعة الدول في الفترة 2000. تذكر أن اختيار برنامج إكسل لتحديد الفئات أن الفئة فيه تتراوح من صفر إلى 2000. وبذلك، يمكننا القول بأن "الفئة من صفر إلى 2000 هى القيمة الأكثر حدوثاً (Modal)". وبدلاً عن ذلك، فإنه من الشائع أخذ القيمة المتوسطة من الفئة المناسبة واعتبارها المنوال. في هذه الحالة، يمكننا القول بأن "المنوال هو 1000 دولار أمريكي". المنوال هو في الغالب أقل شيوعاً في الاستخدام من بين مقاييس النزعة المركزية.

ولفهم الوسيط، تخيل أن كل مساحة المدرج التكراري مظللة. الوسيط هو النقطة على الإحداثي X (x-axis) التى تقسم بدقة هذه المساحة المظللة إلى نصفين. بالنسبة للشكل رقم (2-4) فإن النقطة العليا (بمعنى آخر الوسيط)، هي نقطة المتوسط والتي تقسم الشكل إلى نصفين، وهى في نفس الوقت تعتبر المتوسط الحسابي.

إلا أنه في الشكل رقم (2-2) المتوسط الحسابي هو (5.443.80 دولاراً أمريكياً) الوسيط (3.071.50 دولاراً أمريكياً) والمنوال (1000 دولار أمريكي) وهى مختلفة تماماً.

تعتمد الملخصات الإحصائية المفيدة على مفهوم النسبة المئوية. ومن خلال النظر إلى مجموعة بياناتنا حول متوسط دخل الفرد من إجمالي الناتج المحلى،

وبالنسبة لأى دولة مختارة، فلتكن بلجيكا، يمكنك أن تسأل "كم دولة أكثر فقراً من بلجيكا؟" أو بطريقة أكثر دقة "ماهى نسبة الدول الأكثر فقراً من بلجيكا؟". وعندما نسأل مثل هذه الأسئلة، نحن نسأل بالأحرى ماهى الرتبة النسبية لبلجيكا. من الناحية المنهجية، تمثل الرتبة المئوية لـ "X" قيمة البيانات (مثلا، متوسط نصيب الفرد من إجمالى الناتج المحلى) بحيث إن X% من المشاهدات (مثلا الدول) لها قيمة بيانات أقل. في بيانات إجمالى الناتج المحلى المقطعية تبلغ قيمة الرتبة السابعة والثلاثين 2.092 دولار وهى تمثل متوسط نصيب الفرد من إجمالي الناتج المحلي في بيرو. وهذا يعنى أن 37% من الدول في هذه المجموعة أكثر فقراً من بيرو.

تتعلق العديد من الرتب (Percentiles) بالمفاهيم التى ناقشناها في السابق. فالرتبة 50 (50^{th} percentile) هى الوسيط. وأدنى رتبة وأعلاها هما الصفر 0^{th} والمائة 100^{th}. تقسم الرتبة البيانات إلى مئة درجة، في حين أن المفاهيم الأخرى تستخدم وحدات أساسية أخرى. الربيعات (Quartiles) تقسم مدى البيانات إلى أرباع. وبذلك، يكون الربيع الأول مساويا للرتبة الـ 25، الربيع الثاني يساوي الرتبة الـ 50 (بمعنى آخر الوسيط)، والربيع الثالث الرتبة الـ 75. العشيرات (Deciles) تقسم البيانات إلى عشرات. بمعنى آخر، فإن المدى العشري الأول يساوي الرتبة عشرة، المدى العشري الثاني يساوي الرتبة العشرين .... الخ.

يعتبر المدى الربيعي (Interquartile range) من أكثر قياسات التشتت شيوعاً بعد الانحراف المعياري. وكما يشير اسمه، فإنه يقيس الفرق بين الربيع الثالث والربيع الأول. أما فيما يتعلق بمجموعة البيانات المقطعية، فإن 75% من الدول لديها متوسط دخل فرد من إجمالي الناتج المحلي أقل من 9802 دولار

أمريكي و25% منها لديها متوسط دخل من إجمالي الناتج المحلي أقل من 1162 دولاراً أمريكياً.

بمعنى آخر، فإن 1162 دولاراً أمريكياً هى الربيع الأول، و 9802 دولار أمريكي هو الربيع الثالث. المدى الربيعي هو 9802-1162 = 8640 دولاراً أمريكياً.

ملاحظـــات ختامية:

1- حسب ما تم التأكيد عليه في الفصل الأول، فهذا الكتاب لايُعنى بجمع البيانات، ولكن هذا لا يمنع من إعطاء القارئ بعض الإرشادات حول بعض السبل المثلى للبحث عن البيانات.

2- في بعض الأحيان قد يكون من الملائم أخذ اللوغاريثم الطبيعي أو(-)ln للمتغيرات، وسوف يتم الحديث عن ذلك في الفصول التالية. ويمكن التعرف على اللوغاريثمات وخصائصها الرياضية من خلال أي كتاب يوفر مقدمة في مادة الرياضيات. وباستخدام خصائص اللوغريثمات، يمكن أن نوضح بأن التغير النسبي في المتغير هو $100X[\ l_n(Y_t)-l_n(Y_{t-1})\ ]$. هذه المعادلة يتم استخدامها دائماً في التطبيقات العملية ولها ارتباط وثيق بموضوع السلاسل الزمنية غير المستقرة. (انظر الفصول 9و10).

3- لاحظ أن استخدام كلمة (تكرار) تعني هنا "عدد المشاهدات التي تقع داخل الفئة" وهي تختلف نوعاً ما عن استخدام كلمة (تكرار) في تحليل السلاسل الزمنية (انظر النقاش السابق حول بيانات السلاسل الزمنية).

4- برنامج إكسل يرسم المدرج التكراري مستخدماً مفتاح أمر المدرج التكراري في قائمة (Tools/Data Analysis) وهو يقوم بسهولة برسم المجموعات على المحور الأفقي والتكرار (أو عدد مرات المشاهدات في المجموعة) على المحور الرأسي. لاحظ أن غالبية كتب الإحصاء ترسم الفئات في مقابل التكرارات مقسومة على عرض أو مدى الفئة. هذه الملاحظة الأخيرة تعد صحيحة إذا أُخذ في الاعتبار أن عرض الفئة قد يختلف في جميع الفئات. بمعنى آخر، أن برنامج إكسل لا يقوم بحساب المدرج التكراري بطريقة صحيحة. وعلى الرغم من أن الفئات هي نفس العرض (أو تقريباً كذلك) فإن هذا الخطأ ليس بالأهمية الكبيرة من الناحية العملية.

5- على الطالب الذي يدرس التمويل ولديه اهتمام في متابعة نقاشنا السابق حول متوسط داوجونز الصناعي DJIA، أن ينتبه إلى أن مؤشر S&P500 هو مؤشر مرجح للأسعار يستخدم حجم الشركة كوزن لسعرها في المؤشر.

الفصل الثالث

الارتباط

(Correlation)

يتركز اهتمام الاقتصاديين على اختبار طبيعة العلاقة بين مختلف المتغيرات، مثل مستوى تعليم العاملين وأجورهم أو أسعار الفائدة والتضخم. الارتباط (Correlation) هو طريقة مهمة لقياس وتحديد العلاقة بين متغيرين رقمياً. وهناك مفهوم آخر ذو علاقة وثيقة بالارتباط، سيعرض في الفصول اللاحقة، وهو الانحدار (Regression) الذي يعد امتداداً للارتباط في الحالات التي يكون فيها ثلاثة متغيرات أو أكثر ويقدم فيها مفهوم السببية في العلاقة بين هذه المتغيرات. وسوف تلاحظ سريعاً كلما قرأت في هذا الفصل والفصول التالية، بأن الارتباط والانحدار هما أهم مفهومين يرتبطان ببعضهما بشكل وثيق في هذا الكتاب.

في هذا الفصل، سوف نقوم أولاً بشرح نظرية الارتباط، ثم بعد ذلك نعمل من خلال أمثلة تم تصميمها للمساعدة في التفكير بطريقة بديهية حول مفهوم الارتباط بطرق مختلفة.

فهم الارتباط:

خذ X و Y كمتغيرين (على سبيل المثال، الكثافة السكانية وانحسار مساحة الغابات على التوالي) ودعنا أيضا نفترض بأن لدينا بيانات عن i= 1,.....N وحدات مختلفة (على سبيل المثال، دول). الارتباط بين X و Y يتم تعريفه بالحرف اللاتيني، r، ومعادلته الرياضية موجودة في الملحق رقم (3-1). بالطبع، لن تحتاج إلى استخدام هذه المعادلة بصورة مباشرة؛ لأن الحاسب الآلي يقوم بذلك نيابة عنك. في برنامج إكسل يمكنك استخدام Tools/Data Analysis or Function Wizard لحساب المعادلة الرياضية المشار إليها. وفي العادة يكون واضحاً من خلال السياق لأى المتغيرات ترجع r. إلا أنه في بعض الحالات نكون مضطرين لاستخدام رموز تعريفية (Subscripts) للإشارة إلى أن r_{xy} ترمز للارتباط بين المتغيرين X و Y، r_{xz} ترمز للارتباط بين المتغيرين X و Z ... الخ.

وبمجرد أن تكون قد حسبت معامل الارتباط بين متغيرين سوف تحصل على رقم (على سبيل المثال r= 0.55). ومن المهم أن تعرف كيف تفسر هذا الرقم. في هذا القسم، سوف نحاول استيعاب المنطق حول مفهوم الارتباط، ولكن بداية دعنا نذكر باختصار بعض الخصائص الرقمية للارتباط.

خصائص الارتباط:

1- قيمة r تقع دائما بين -1 و 1 والتي يمكن كتابتها على النحو التالي $(-1 \leq r \leq 1)$.

2- القيم الموجبة للحرف r تشير إلى ارتباط موجب بين X و Y. القيم السالبة تشير إلى ارتباط سالب. 0= r تشير إلى أنه ليس هناك ارتباط بين X و Y.

3- قيم r الموجبة الكبيرة تشير إلى ارتباط موجب قوي. 1= r تشير إلى وجود ارتباط موجب تام. قيم r السالبة الكبيرة[1] تشير إلى ارتباط سالب قوي. -1= r تشير إلى وجود ارتباط سالب تام.

4- الارتباط بين Y و X هو نفس الارتباط بين X و Y.

5- الارتباط بين أي متغير ونفسه (على سبيل المثال، الارتباط بين Y و Y) يساوي 1.

فهم الارتباط من خلال التبرير الكلامي:

يستخدم الإحصائيون مفهوم الارتباط بكثرة وبنفس الطريقة التي يستخدمها الشخص العادي. العرض التالي لمثال انحسار مساحة الغابات/الكثافة السكانية المأخوذ من الفصل الثاني سوف يستفاد منه لشرح الطرق الكلامية لاستيعاب مفهوم الارتباط.

مثال: الارتباط بين معدل انحسار مساحة الغابات والكثافة السكانية.

دعنا نفترض أننا نرغب في اختبار العلاقة بين معدل انحسار مساحة الغابات والكثافة السكانية. تذكر أن ملف إكسل FOREST.XLS يحتوي على بيانات حول هذه المتغيرات (ومتغيرات أخرى) لبيانات مقطعية لـ 70 دولة استوائية. وباستخدام برنامج إكسل، وجدنا أن هناك ارتباطاً بين الانحسار في مساحة الغابات (Y) والكثافة السكانية (X) بمقدار 0.66. وما دام أن هذا الرقم أكبر من الصفر، فإن هذا الرقم يتيح لنا كتابة العبارات التالية:

1- هناك علاقة موجبة (أو ارتباط موجب) بين الانحسار في مساحة الغابات والكثافة السكانية.

2- الدول ذات الكثافة السكانية العالية تميل إلى أن يكون لديها معدلات عالية للانحسار في مساحة الغابات. الدول ذات الكثافة السكانية المنخفضة تميل إلى أن تكون لديها معدلات منخفضة للانحسار في مساحة الغابات. لاحظ أننا استخدمنا هنا عبارة (تميل إلى أن يكون). الارتباط الموجب لايعني بالضرورة أن يكون لكل دولة ذات كثافة سكانية عالية معدل عال للانحسار في مساحة الغابات، إلا أن هذا هو الاتجاه العام للبيانات. ومن المحتمل أن تتبع دول قليلة هذا النمط (انظر النقاش حول الحالات الشاذة في الفصل الثاني).

3- تختلف معدلات الانحسار في مساحة الغابات عبر الدول كما تختلف الكثافات السكانية (ولهذا السبب نسميها بـ "المتغيرات") بعض الدول لديها معدلات عالية للانحسار في مساحة الغابات، وبعضها لديه معدلات منخفضة. هذا التباين بين الدول من زيادة/انخفاض في معدلات الانحسار في مساحة الغابات يميل إلى "مساواة" التباين من زيادة/ انخفاض في الكثافات السكانية.

كل العبارات السابقة تتطلب أن تكون قيمة r موجبة. فإذا كانت قيم r سالبة فإن عكس العبارات السابقة يكون صحيحاً. على سبيل المثال قيم X العالية تكون مرتبطة بقيم منخفضة لـ Y ... الخ.

وجدير بالذكر أنه يصعب أن تشعر حسياً بالمدلول الدقيق لرقم معامل الارتباط (على سبيل المثال، كيف يختلف معامل الارتباط 0.66 عن معامل الارتباط 0.26). العلاقة بين المتغيرين x و y (XY-plots) التي سيتم مناقشتها أدناه تقدم بعض المساعدة، إلا أننا سوف نؤكد هنا باختصار نقطة مهمة وسنرجع إليها عندما نناقش موضوع الانحدار.

4- الدرجة التي تختلف بها معدلات الانحسار في مساحة الغابات عبر الدول يمكن قياسها رقمياً باستخدام معادلة الانحراف المعياري التي تمت مناقشتها في الفصل الثاني. وكما ورد في النقطة 3 أعلاه، فإن حقيقة أن الانحسار في مساحة الغابات والكثافة السكانية لهما ارتباط موجب تعني أن أنماط تغيرهما بين الدول تميل إلى أن تكون متساوية. تربيع معامل الارتباط (r^2) يقيس نسبة التغير بين الدول في الانحسار في مساحة الغابات التي يساويها، أو يفسرها، التباين في الكثافة السكانية. بمعنى آخر، معامل الارتباط هو مقياس رقمي للدرجة التي تتقابل عندها أنماط كل من X وY. في مثالنا عن الكثافة السكانية / الانحسار في مساحة الغابات، فطالما أن ($0.66^2 = 0.44$)، يمكننا القول أن 44% من التباين بين الدول في معدل الانحسار في مساحة الغابات يمكن تفسيره من خلال التباين في الكثافة السكانية.

تمرين رقم (3-1):

(أ) باستخدام البيانات في الملف FOREST.XLS، احسب، فسر المتوسط الحسابي، الانحراف المعياري، أقل وأكثر انحسار في مساحة الغابات والكثافة السكانية.

(ب) أثبت أن معامل الارتباط بين المتغيرين هو 0.66.

مثال: أسعار المنازل في مدينة ويندسور بكندا

يشتمل ملف إكسل HPRICE.XLS على بيانات 546 منزلاً تم بيعها في مدينة ويندسور بكندا خلال صيف عام 1987م. وتحتوي تلك البيانات على سعر البيع (بالدولار الكندي) إضافة للعديد من المواصفات المتعلقة بكل منزل. وسوف نستخدم هذه البيانات بصورة مكثفة في الفصول اللاحقة، وفي الوقت الحاضر دعنا نركز على متغيرات قليلة فقط، وتحديداً، دعنا نفترض أن Y تساوي سعر بيع المنزل و X تساوي مساحته محسوبة بالأقدام المربعة، وحجم المنزل هو المساحة التي يشغلها المنزل نفسه إضافة لحديقته أو المساحة الخالية منه.

الارتباط بين هذين المتغيرين هو $r_{xy} = 0.54$

يمكن صياغة العبارات التالية حول أسعار المنازل في ويندسور:

1- المنازل ذات المساحات الكبيرة تكون أعلى سعراً مقارنة بتلك المنازل ذات المساحات الصغيرة.

2- هناك علاقة موجبة بين مساحة المنزل وسعر المنزل.

3- تسهم مساحة المنازل بمعدل 29% من التباين في أسعارها (بمعنى آخر $0.54^2 = 0.29$).

الآن دعنا نضيف متغيراً ثالثاً، Z، ويمثل عدد غرف النوم.

وبحساب معامل الارتباط بين أسعار المنازل وعدد غرف النوم، نحصل على $r_{yz} = 0.37$. توضح هذه النتيجة، كما نتوقع، إلى أن المنازل ذات العدد الأكثر من غرف النوم تكون ذات سعر أعلى من المنازل ذات عدد غرف النوم الأقل.

وبنفس الأسلوب، يمكننا حساب معامل الارتباط بين عدد غرف النوم ومساحة المنزل. هذا الارتباط يساوي $r_{xz} = 0.15$ ويشير إلى أن المنازل ذات المساحة الأكبر تكون بها غرف نوم أكثر. وعلى الرغم من ذلك، فإن معامل الارتباط يُعد صغيراً جداً، وغير متوقع إطلاقاً، وربما يرجع السبب في ذلك إلى أن العلاقة بين مساحة المنزل وعدد غرف النوم ضعيفة جداً. بمعنى آخر، قد تتوقع أن المنازل المشيدة على مساحات كبيرة، يكون بها عدد أكثر من غرف النوم مقارنة بالمنازل المشيدة على مساحات صغيرة. إلا أن معامل الارتباط يشير إلى أن هناك احتمالاً ضعيفاً لحدوث ذلك.

المثال أعلاه يتيح لنا أن نثير هنا قضية مهمة في علم الاقتصاد القياسي وهي قضية السببية (Cavsality)، وذلك لأن الاقتصاديين يهتمون بمعرفة ما إذا كان متغير ما (يتسبب) في حصول تغيير في متغير آخر. ولن نقدم هنا تعريفاً نظرياً للسببية وبدلاً عن ذلك سوف نستخدم الكلمة في معناها المتداول يومياً. في هذا المثال، قد يكون من المفيد استخدام الارتباط الموجب بين سعر المنزل ومساحة الأرض ليعكس العلاقة السببية. أي أن مساحة الأرض هي المتغير التي تؤثر مباشرة (أو تتسبب) في أسعار المنازل. على الرغم من أن أسعار المنازل لا

تؤثر (أو تتسبب) في المساحة. بمعنى آخر، أن اتجاه السببية يكون من مساحة الأرض إلى أسعار المنازل، وليس العكس.

هناك طريقة أخرى للتفكير في هذه القضايا تتمثل في أن تسأل نفسك عن ما سيحدث إذا رغب مالك المنزل شراء أرض مجاورة لمنزله، ونتيجة لذلك زادت مساحة منزله. هذا العمل سوف يؤدي إلى زيادة قيمة المنزل (بمعنى آخر- أن الزيادة في مساحة الأرض سوف تتسبب في ارتفاع سعر المنزل). إلا أنه إذا نظرت إلى سؤال عكس ذلك مثل: "هل تتسبب زيادة سعر المنزل في زيادة مساحة الأرض؟" سوف تجد أن السببية العكسية غير موجودة (بمعنى آخر، ارتفاع سعر المنزل لا يتسبب في زيادة مساحة الأرض). على سبيل المثال، إذا زادت أسعار المنازل في ويندسور بطريقة مفاجئة لسبب ما (على سبيل المثال، بسبب حدوث ازدهار اقتصادي) فإن هذا لا يعني أن المنازل في ويندسور قد حصلت فجأة على مساحات أكبر.

النقاش الذي أجريناه في الفقرة السابقة يمكن تكراره من خلال إبدال (مساحة الأرض) (بعدد غرف النوم). أي أنه من المناسب الافتراض بأن الارتباط الموجب بين Y تمثل أسعار المنازل و Z تمثل عدد غرف النوم وذلك بسبب أن Z تؤثر (أو تتسبب) في Y، وليس العكس. لاحظ، مع ذلك، أنه من الصعوبة توضيح الارتباط الإيجابي (إلا أنه ضعيف) بين X التي تساوي مساحة الأرض و Y تمثل عدد غرف النوم بحسب السببية الموضحة. أي أن هناك ميلاً لأن تشغل المنازل ذات غرف النوم الأكثر مساحات كبيرة، إلا أن هذا الميل لا يعني أن عدد غرف النوم هو الذي يتسبب في المساحات الكبيرة للمنازل.

تعد المعرفة بكيفية تفسير النتائج التي تتوصل إليها من أهم الصعوبات في العمل التطبيقي (القياسي). مثال المنزل الوارد أعلاه، يوضح هذه الصعوبة بجلاء. ولا يكفي إظهار رقم الارتباط فقط (على سبيل المثال $r_{xy} = 0.54$)،

فالتفسير والتعليل مهم أيضاً. يتطلب التفسير معرفة جيدة عن ماهية الارتباط بالإضافة إلى قدر كبير من الفهم العميق تجاه الظاهرة الاقتصادية محل الدراسة. وبالأخذ في الاعتبار أهمية التفسير في البحث القياسي، فإن القسم التالي سوف يقدم العديد من الأمثلة لإظهار لماذا تكون المتغيرات ذات ارتباط ببعضها وكيف يمكن أن يقودنا الفهم البديهي إلى توضيحها.

تمرين رقم (3-2):

(أ) باستخدام البيانات في الملف HPRICE.XLS، احسب واشرح المتوسط الحسابي، الانحراف المعياري، أدنى وأعلى قيمة للمتغيرات Z,Y,X حيث تمثل Y سعر المنزل (المعنونة "سعر البيع" في الملف HPRICE.XLS)، وتمثل X مساحة الأرض وتمثل Z عدد غرف النوم (المعنونة "bedroom").

(ب) اثبت أن الارتباط بين X وY هو مثل ما هو معطى في المثال. كرر الإثبات بالنسبة لكل من X وZ وبعد ذلك بالنسبة لكل من YوZ.

(ج) الآن أضف متغيراً جديداً، W تمثل عدد الحمامات (معنونة "bath") احسب المتوسط الحسابي لـ W.

(د) احسب وفسر الارتباط بين W و Y. ناقش إلى أي مدى يمكن القول بأن W تسبب Y.

(هـ) كرر الفقرة (د) بالنسبة لكل من W و X وبعد ذلك بالنسبة لكل من Wو Z.

معرفة لماذا تكون المتغيرات في حالة ارتباط:

في مثالنا عن الانحسار في مساحة الغابات/ الكثافة السكانية، اكتشفنا أن الانحسار في مساحة الغابات والكثافة السكانية هما في الواقع في حالة ارتباط إيجابي، مما يشير إلى وجود علاقة إيجابية بين الاثنين. ولكن ما هو الشكل الذي تأخذه هذه العلاقة؟ كما تم مناقشته أعلاه، فإننا غالباً نرغب في التفكير بصيغة السببية والتأثير، وقد تكون في الواقع هي حالة تشير إلى أن الارتباط والسببية يرتبطان ببعضهما بشكل وثيق.

على سبيل المثال، فإن التوصل إلى أن الكثافة السكانية والانحسار في مساحة الغابات في حالة ارتباط قد يعني أن الكثافة السكانية تتسبب بشكل مباشر في حدوث الانحسار في مساحة الغابات. وبالمثل، فإن التوصل إلى وجود ارتباط موجب بين مستويات التعليم والأجور يمكن توضيحه بما يعني أن المزيد من التعليم يؤثر بشكل مباشر في الأجر الذي يتقاضاه الفرد. ومع ذلك، فإن الأمثلة التالية توضح حقيقة أن الارتباط يعني أن السببية لا تكون بالضرورة دائماً صحيحة.

مثال: الارتباط لايعني بالضرورة السببية:

من المسلم به على نطاق واسع أن التدخين يسبب سرطان الرئة. دعنا نفترض أننا جمعنا بيانات عن العديد من الأشخاص حول:

(أ) عدد السجائر التي يدخنها كل شخص في الأسبوع (X).

(ب) ما إذا كانوا قد أصيبوا من قبل أو يعانون الآن من سرطان الرئة (Y).

وما دام التدخين يسبب السرطان فإننا بدون شك سوف نجد أن $r_{xy} > 0$، أي أن الأفراد الذين يدخنون سوف تكون معدلات إصابتهم بسرطان الرئة أعلى من غير المدخنين. وهنا يشير الارتباط الموجب بين X وY إلى وجود علاقة سببية مباشرة.

الآن افترض بأن لدينا بيانات حول نفس الأشخاص، تقيس الكمية التي يشربونها من المشروبات الكحولية في نفس الأسبوع. ودعنا نسمي هذا المتغير بالحرف Z. في العادة يكون الأشخاص الذين يشربون الكحول بكثافة من المدخنين، لذلك، فإن $r_{xz} > 0$. هذا الارتباط لا يعني أن تدخين السيجارة لايتسبب في أن يجعل الأفراد يشربون الكحول. وإنما من المحتمل أن يعكس ذلك بعض السلوكيات الاجتماعية الدارجة. بمعنى آخر، ربما تعكس حقيقة أن الأفراد الذين يدخنون لا يقلقون على صحتهم أو أن حياتهم الاجتماعية تدور حول الحانات، حيث يكون شرب الكحول وتدخين السجائر مترافقان في أغلب الأحيان. وفي كل حالة، فإن الارتباط الموجب بين التدخين وشرب الكحول قد يعكس سبباً ما (على سبيل المثال، سلوك اجتماعي) والذي بدوره يتسبب في الاثنين معاً. لذلك، فإن الارتباط بين المتغيرين لا يعني بالضرورة أن أحدهما يسبب الآخر. وقد تكون حالة مسئول عنها متغير ثالث غير واضح.

الآن انظر للارتباط بين سرطان الرئة وإدمان شرب الكحول. وما دام الأشخاص الذين يدخنون يكونون أكثر قابلية للإصابة بسرطان الرئة، وأن الأشخاص الذين يدخنون يميلون لشرب الكحول أكثر من غيرهم، فإنه من غير المعقول أن تتوقع بأن معدلات سرطان الرئة سوف تكون عالية بين من يدمنون شرب الكحول (أي أن $r_{yz} > 0$). لاحظ أن هذا الارتباط الموجب لا يعني أن استهلاك الكحول يسبب سرطان الرئة. وبدلاً عن ذلك فإن تدخين السجائر هو الذي يسبب السرطان، إلا أن التدخين وشرب الكحول يتعلقان بسلوك اجتماعي معين. هذا المثال يشير إلى نوع من الأنماط المعقدة للسببية التي تحدث في الواقع، وكيف يجب أخذ الحذر عند محاولة ربط مفاهيم الارتباط والسببية.

مثال: السببية المباشرة في مقابل السببية غير المباشرة:

يعد الفرق بين السببية المباشرة والسببية غير المباشرة فرقاً مهماً. واستدراكاً لمثلنا عن الانحسار في مساحة الغابات/ الكثافة السكانية، حيث وجدنا أن الكثافة السكانية (X) والانحسار في مساحة الغابات (Y) هما في حالة ارتباط موجبة (أي أن $r_{xy} > 0$). هناك سبب لهذا الارتباط الموجب يتمثل في أن الضغوطات السكانية في المناطق الريفية تدفع المزارعين لقطع الغابات لتجهيز أرض جديدة من أجل زراعة محاصيل زراعية، أي أن عملية التوسع الزراعي المستمرة هي التي تسبب الانحسار في مساحة الغابات بصورة مباشرة. إذا حسبنا الارتباط بين الانحسار في مساحة الغابات ومساحة الأراضي الزراعية (Z)، فسوف نجد أنه $r_{yz} > 0$. وفي هذه الحالة تكون الكثافة السكانية هي السبب غير المباشر، والتوسع الزراعي هو السبب المباشر لحدوث الانحسار في مساحة الغابات. بمعنى آخر، يمكننا القول أن X (الزيادة في عدد السكان) تسبب Z (التوسع الزراعي) والتي تسبب في المقابل Y (الانحسار في مساحة الغابات). مثل هذا النمط من السببية يتسق مع $r_{xy} > 0$ و $r_{zy} > 0$.

إلا أنه في مثالنا عن سعر المنزل، فإنه من المتوقع أن تعكس الارتباطات الموجبة التي لاحظناها السببية المباشرة. على سبيل المثال، امتلاك قطعة أرض كبيرة يعتبرها معظم الأفراد ميزة جيدة في ذاتها، لذلك فإن زيادة مساحة الأرض يجب أن تؤدي مباشرة إلى زيادة سعر المنزل. وليس هناك متغير آخر يؤثر في هذا الموضوع، ولذلك فإننا نقول إن السببية مباشرة[2].

الاستفادة العامة التي يجب أن نأخذها من هذه الأمثلة هي أن الارتباطات قد تكون موجبة جداً، إلا أنها قد لا تنشئ في حد ذاتها سببية. في مثال التدخين/ السرطان أعلاه، فإن اكتشاف الارتباط الموجب بين التدخين وسرطان الرئة،

المرتبط بالإثبات الطبي عن الحالة التي تحفز بها المواد الموجودة في السجائر التغيرات في الجسم البشري مما يؤثر في معظم المدخنين متسببة في السرطان. في مثال سعر المنزل، نعرف بالفطرة أن متغير عدد غرف النوم يؤثر بشكل مباشر في أسعار المنازل. في علم الاقتصاد، يمكن استخدام مفهوم الارتباط، بالارتباط مع الإحساس البديهي أو النظرية الاقتصادية المقنعة لحدوث السببية.

تمرين رقم (3-3):

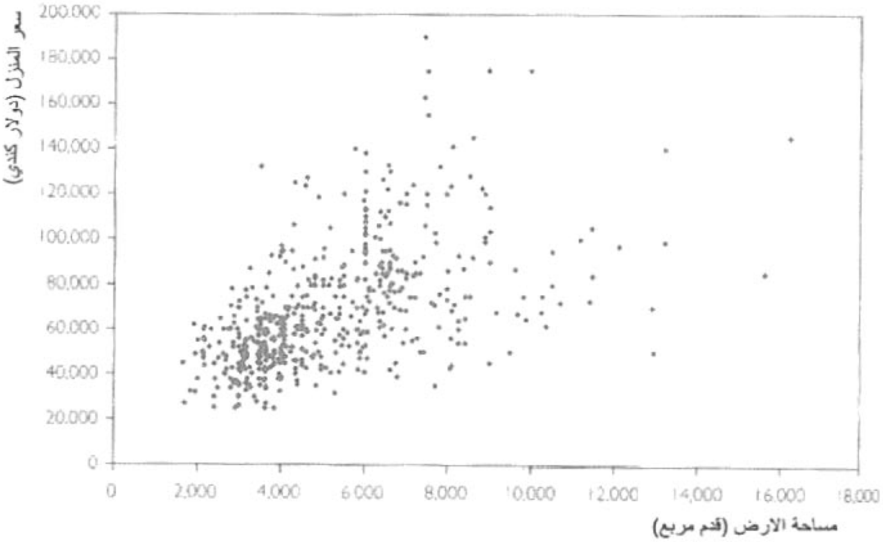
الأشخاص الذين نالوا تعليماً جامعياً يميلون إلى شغل وظائف ذات رواتب أعلى من أولئك الذين لديهم مؤهلات تعليمية أقل. وقد يُعزى ذلك لحقيقة أن التعليم الجامعي يوفر مهارات مهمة تعود على أصحاب العمل بقيمة عالية. أو قد يكون الوضع الذي يجعل الأشخاص الأذكياء يذهبون للجامعة، وأن أصحاب العمل يرغبون في تشغيل هؤلاء الأذكياء (أي أن الدرجة الجامعية ليست محل اهتمام أصحاب العمل في ذاتها).

افترض أن لديك بيانات عن Y تمثل الدخل، X تمثل عدد سنوات الدراسة و Z تمثل نتائج اختبار الذكاء [3] للعديد من الأشخاص، وأنك حسبت قيمة كل من r_{xz} ، r_{xy} و r_{yz} ، في الواقع، ماهي الإشارات التي تتوقع أن تحملها هذه الارتباطات؟ وبافتراض أن الارتباطات تحمل الإشارات التي تتوقعها، هل يمكنك تحديد أي من المثالين في الفقرة أعلاه يعتبر صحيحاً؟

فهم الارتباط من خلال شكل الانتشار:

يمكن استيعاب مدلولات الارتباط من خلال شكل الانتشار بين متغيرين الذي تم إيضاحه في الفصل الثاني، آخذين في الاعتبار أننا سوف نناقش في هذا الفصل العلاقات الموجبة والسالبة المبنية على ما إذا كان شكل الانتشار يوضح ميلاً عاماً إلى أعلى أو إلى أسفل[4]. إذا كان هناك ارتباط بين متغيرين فإن شكل الانتشار لأحدهما في مقابل الآخر سوف يوضح أيضاً هذه الأنماط. على سبيل المثال يعرض شكل الانتشار للكثافة السكانية في مقابل الانحسار في مساحة الغابات نمط ميل إلى أعلى، (انظر الشكل رقم 3-2). يدّل هذا الانتشار على وجود حالة ارتباط إيجابي بين المتغيرين، ووجدنا أن ذلك هو الواقع من خلال معامل الارتباط $r = 0.66$. النقطة المهمة هنا هي أن الارتباط الموجب يرتبط بأنماط الميل إلى أعلى في شكل الانتشار والارتباط السالب يرتبط بأنماط الميل إلى أسفل. كل الفهم البديهي الذي طورناه حول أشكال الانتشار في الفصل السابق يمكن الآن استخدامه لتطوير فهم أعمق حول الارتباط.

يستخدم الشكل رقم (3-1) بيانات أسعار منازل (House Price) ويندسور المتوافرة في الملف (HPRICE.XLS) لرسم شكل الانتشار، حيث تمثل X مساحة الأرض (Lot Size) في مقابل Y التي تمثل سعر المنزل، آخذين في الاعتبار أن الارتباط بين هذين المتغيرين قد تم حسابه، وهو يساوي $r_{xy} = 0.54$، وهو رقم موجب.



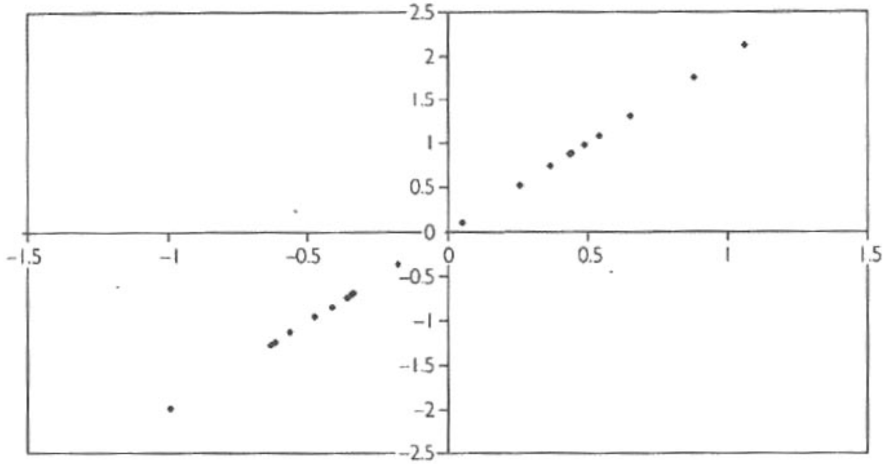
شكل رقم (3-1) سعر المنزل مقابل مساحة المنزل

هذه العلاقة الإيجابية (الميل إلى أعلى) بين مساحة الأرض وسعر المنزل يمكن رؤيتها بوضوح في الشكل رقم (3-1)، إذ تشير إلى أن المنازل ذات المساحات الصغيرة (بمعنى آخر قيم صغيرة على محور X) تميل إلى أن يكون لها أسعار منخفضة (بمعنى آخر، قيم أقل على محور Y). وعلى العكس من ذلك، فإن المنازل ذات المساحات الكبيرة يكون لها أسعار عالية.

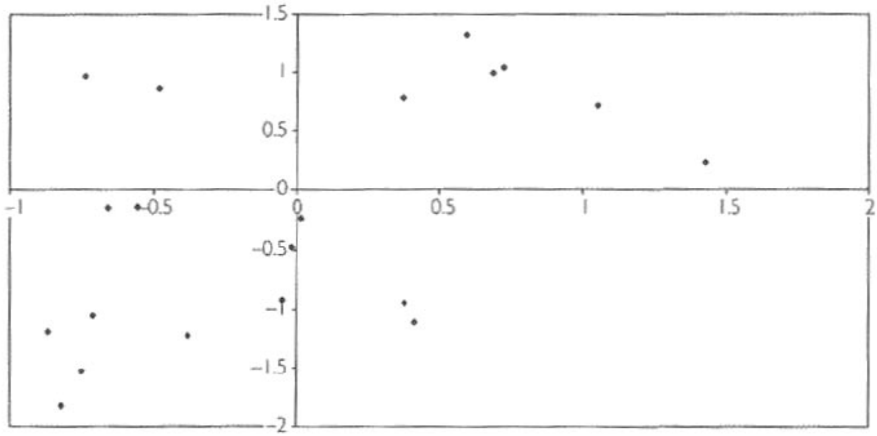
النقاش السابق يتعلق أساساً بإشارة الارتباط. إلا أنه يمكن أيضاً استخدام شكل الانتشار لتطوير فهمنا لتفسير حجم رقم الارتباط، كما توضحه الأمثلة التالية.

الشكل رقم (3-2) هو عبارة عن شكل الانتشار لمتغيرين في حالة ارتباط كامل (بمعنى آخر أن $r = 1$). لاحظ أن بيانات هذين المتغيرين ليست بيانات اقتصادية حقيقية، وإنما تم استحداثها من خلال الحاسب الآلي. لاحظ أن كل النقاط تقع بالضبط على خط مستقيم.

الشكل رقم (3-3) هو عبارة عن شكل الانتشار لمتغيرين في حالة ارتباط موجب ($r = 0.51$)، إلا أنهما ليسا في حالة ارتباط كامل. لاحظ أن شكل الانتشار مازال يشير إلى نمط ميل إلى أعلى، إلا أن النقاط تنتشر بشكل أكثر اتساعاً.



شكل رقم (3-2) شكل الانتشار لمتغيرين في حالة ارتباط كامل ($r = 1$)



شكل رقم (3-3) شكل الانتشار لمتغيرين في حالة ارتباط موجب ($r = 0.51$)

الشكل رقم (3-4) هو عبارة عن شكل الانتشار لمتغيرين في حالة انعدام الارتباط بالكامل ($r = 0$). لاحظ أن النقاط منتشرة بشكل عشوائي على كامل مساحة الانتشار. تشير أشكال الانتشار للارتباط السالب إلى أنماط ميل إلى أسفل، إلا أن عكس ذلك هو انطباق نفس أنواع الأنماط المشار إليها أعلاه عليهما. وعلى سبيل المثال، الشكل رقم (3-5) هو شكل الانتشار لمتغيرين في حالة ارتباط سالب (r=-0.58)

تعبر هذه الأشكال (أشكال الانتشار) إلى أحد الطرق للتفكير بمفهوم الارتباط، الارتباط يشير إلى كيف يمكن رسم خط مستقيم بطريقة جيدة بحيث يتوسط نقاط الانتشار، وبحيث تقع المتغيرات ذات الارتباط القوي ببعضها على هذا الخط المستقيم أو بالقرب منه، أما المتغيرات ذات الارتباط الضعيف فتكون أكثر انتشاراً على شكل الانتشار.

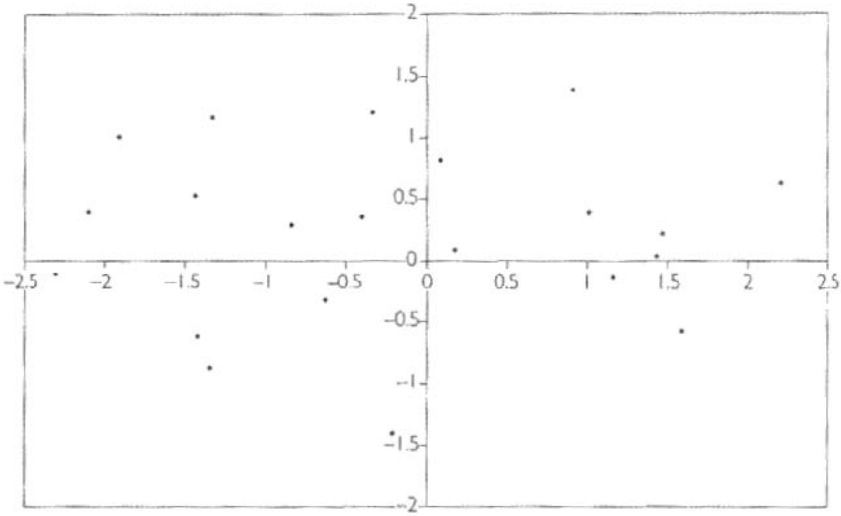
تمرين رقم (3-4):

يشتمل الملف EX34.XLS على أربعة متغيرات Y، X_1، X_2، وX_3

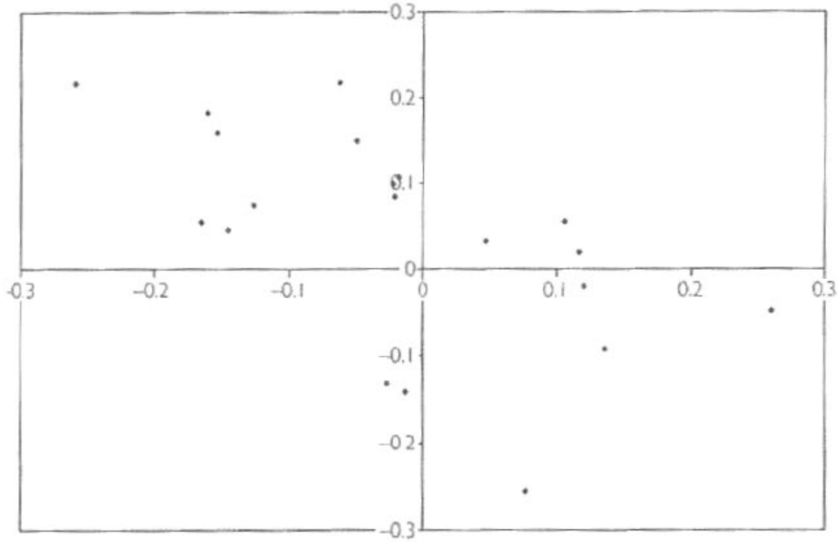
(أ) احسب الارتباط بين YوX_1. كرر نفس العملية لـ X_2 وY و لـX_3 و Y.

(ب) ارسم شكل الانتشار للمتغيرين Y و X_1 . كرر نفس العملية لـ Y و X_2 ولـ Y و X_3.

(جـ) فسر النتائج التي توصلت إليها في الفقرتين (أ) و (ب).



شكل رقم (3-4) انتشار متغيرين في حالة عدم ارتباط (r=0)



شكل رقم (3-5) انتشار متغيرين في حالة ارتباط سالب (r=-0.58)

حالة الارتباط بين متغيرات متعددة:

الارتباط هو خاصية تربط بين متغيرين مع بعضهما، إلا أنه يُتطلب من الاقتصاديين باستمرار العمل على متغيرات متعددة. على سبيل المثال، تعتمد أسعار المنازل على مساحة الأرض، عدد غرف النوم، عدد الحمامات وعلى العديد من مواصفات المنزل. وكما سوف نرى في الفصول اللاحقة، أن الانحدار هو الأداة المناسبة للاستخدام إذا كان التحليل يشتمل على أكثر من متغيرين. ومع ذلك فإنه من غير المعتاد للباحثين القياسيين - عندما يعملون على متغيرات متعددة - حساب الارتباط بين كل زوجين من المتغيرات. هذا الحساب يأخذ جهداً كبيراً عندما يكون عدد المتغيرات كبيراً. على سبيل المثال، إذا كان لدينا ثلاثة متغيراتX,Y, وZ فإن هناك ثلاثة ارتباطات محتملة (بمعنى آخر، r_{xy}، r_{xz}، r_{yz}).

إلا أنه إذا أضفنا متغيراً رابعاً، W، فإن عدد الارتباطات سوف يرتفع إلى ستة (بمعنى آخر r_{xy}، r_{xz}، r_{xw}، r_{yz}، و r_{zw}. وعموماً، فإنه بالنسبة لمتغيرات مختلفة M فإن الارتباطات المحتملة هي $M\times(M-1)/2$.

الطريقة المناسبة لوضع كل هذه الارتباطات في إطارها الصحيح يتمثل في إنشاء مصفوفة أو جدول، كما يوضح ذلك المثال التالي.

يشتمل ملف CORMAT.xls على بيانات عن ثلاثة متغيرات هي Z,Y,X. المتغير X في العمود الأول، المتغير Y في العمود الثاني والمتغير Z في العمود الثالث. وباستخدام برنامج إكسل يمكننا إنشاء مصفوفة ارتباط (الجدول 3-1) لهذه المتغيرات.

الرقم 0.318237 هو الارتباط بين المتغير في العمود الأول (X)، والمتغير في العمود الثاني (Y). وبنفس الطريقة فإن الرقم ‎-0.13097‎ هو الارتباط بين Z و X والرقم 0.096996 هو الارتباط بين Z,Y.

لاحظ أن أرقام الواحد الصحيح تشير إلى أن كل متغير هو في حالة ارتباط تام مع نفسه.

جدول رقم (3-1) مصفوفة الارتباط لكل من Y,X وZ

	العمود 1 (Column 1) (X)	العمود 2 (Column 2) (Y)	العمود 3 (Column 3) (Z)
عمود 1 Column 1(X)	1	———	———
عمود 2 Column 2(Y)	0.318237	1	———
عمود 3 Column 3(Z)	- 0.13097	0.096996	1

تمرين رقم (3-5):

(أ) مستخدما البيانات في الملف FOREST.XLS، احسب وفسر مصفوفة الارتباط متضمنة معدل الانحسار في مساحة الغابات، الكثافة السكانية، التغير في مساحة المراعي والتغير في مساحة أراضي زراعة المحاصيل الزراعية.

(ب) كرر الجزء (أ) مستخدماً المتغيرات في ملف البيانات HPRICE.XLS: سعر المنزل، مساحة الأرض، عدد غرف النوم، عدد الحمامات وعدد الأدوار. كم عدد الارتباطات التي حسبتها

ملخص الفصـــل:

1- الارتباط هو طريقة شائعة لقياس العلاقة بين متغيرين. وهو رقم يمكن حسابه باستخدام برنامج إكسل أو برامج حاسب آلي تطبيقية أخرى.

2- الارتباط يمكن توضيحه بطريقة بديهية كقياس رقمي لعلاقة بين متغيرين.

3- الارتباط يمكن أيضا توضيحه بالرسم باستخدام شكل الانتشار بين متغيرين. أي أن علامة الارتباط تتعلق أو ترتبط بميل أفضل خط اتصال خلال شكل الانتشار. حجم الارتباط يتعلق بكيفية انتشار نقاط البيانات حول أفضل خط التقاء.

4- تتكون الارتباطات بين متغيرين لعدد من الأسباب. ومع ذلك، فإن الارتباط لا يعنى بالضرورة وجود سببية بين متغيرين.

ملحق رقم (3-1): التفاصيل الرياضية:

الارتباط بين X وY يرمز إليه بحرف صغير r ويتم حسابه بالمعادلة:

$$r = \frac{\sum_{i=1}^{N}(Y - \bar{Y})(X_i - \bar{X})}{\sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2}}$$

حيث $\bar{X}$ و $\bar{Y}$ هما المتوسط الحسابي لكل من X وY (انظر الفصل الثاني). وبطريقة أكثر بديهية، لاحظ أننا إذا قسمنا البسط والمقام للمعادلة السابقة على N-1 فسوف يحتوي المقام على ناتج ضرب (مضروب) الانحراف المعياري لكل من Xو Y، والبسط على التغاير بين كل من YوX. التغاير (Covariance) هو مفهوم لم نقم بتعريفه هنا، ولكنه سوف يقابلك في المستقبل، وبخاصة إذا كنت مهتماً بالتوسع في فهم النظرية الإحصائية التي تتضمن الارتباط.

ملاحظـــات ختامية:

1- نعني بعبارة "القيم السالبة الأكبر" الأكثر سلبية. على سبيل المثال، -0.9 أكثر سلبية من قيمة -0.2.

2- التوضيح البديل هو أن الأحياء الراقية فيها منازل ذات مساحات كبيرة. الأفراد مستعدون لدفع مبالغ أكبر من أجل أن يعيشوا في حي سكني راق. لذلك، فإنه من الممكن أن يكون للمنازل ذات المساحات الكبيرة أسعار عالية، ليس لأن الأفراد يرغبون في المساحات الكبيرة، بقدر رغبتهم في العيش في حي سكنى راق. بمعنى آخر، قد تعمل "مساحة الأرض" كمرجعية لأثر "الحى السكنى الراقي". سوف نناقش مثل هذه القضايا بتفصيل أكثر في الفصول اللاحقة عن الانحدار. يجب أن تعلم تماماً هنا أن توضيح الارتباطات قد يكون معقداً جداً وقد يتسق نمط الارتباط المعطى مع أمثلة وشروحات بديلة كثيرة.

3- تعد قضية ما إذا كانت اختبارات الذكاء مقاييس حقيقية وذات معنى للذكاء، إحدى القضايا التي تثير الجدل بين علماء النفس والباحثين النفسيين. وللإجابة عن هذا السؤال، تجنبوا هذه الجدلية وافترضوا أن هذه الاختبارات هي انعكاس دقيق للذكاء.

4- سوف نقوم بالتعبير عن معاني أنماط الميل "إلى أعلى" أو "إلى أسفل" في شكل الانتشار عندما نأتي للحديث عن الانحدار. وللمساعدة في التوضيح، فكر في رسم خط مستقيم خلال النقاط في شكل الانتشار يحتوي النمط الموجود في البيانات بشكل جيد (بمعنى آخر، أحسن خط يجمع كل النقاط). حيث يمكن التعبير عن الميل إلى أعلى والميل إلى أسفل بالإشارة إلى ميل هذا الخط المستقيم.

الفصل الرابع

مقدمة إلى الانحدار البسيط

(An Introduction to Simple Regression)

يعد الانحدار البسيط أهم أداة يستخدمها الاقتصاديون في الدراسات القياسية لتحديد وفهم العلاقة بين اثنين أو أكثر من المتغيرات. وهو مفيد خصوصاً في الحالة العامة التي يوجد فيها العديد من المتغيرات (على سبيل المثال، البطالة وأسعار الفائدة، عرض النقود، أسعار الصرف، التضخم ... الخ) وكذلك تكون العلاقات بينها معقدة.

ولإعطاء مثال على ذلك، نشير إلى أن الاهتمام الكبير بوسائل الإعلام في المملكة المتحدة في صيف عام 1988م تركز على المستوى المناسب الذي يجب أن تكون عليه أسعار الفائدة، خاصة بعد شكوى القطاع الصناعي من الارتفاع الحاد لأسعار الفائدة. حيث يشيرون إلى أن أسعار الفائدة العالية تشجع الأجانب على استثمار أموالهم في المملكة المتحدة، ويتسبب ذلك في المقابل في ارتفاع قيمة الجنيه الإسترليني. القيمة المرتفعة للجنيه الإسترليني تجعل من الصعب على مؤسسات الأعمال في المملكة المتحدة تصدير منتجاتها، مما ينتج عنه انخفاض في المبيعات، وزيادة في عدد العمال الذين يفقدون عملهم مما يؤدي إلى ارتفاع معدل البطالة.

وفي مقابل من يعتقدون بارتفاع أسعار الفائدة، هناك آخرون يعتقدون بأن أسعار الفائدة منخفضة جداً، ويقولون بأن أسعار الفائدة المرتفعة تعتبر ضرورية للتخلص من الضغوط التضخمية بسبب العلاقة بين التضخم وأسعار الفائدة. لذلك فإن هذه القضية الاقتصادية المهمة (تحديد سعر الفائدة) ما زالت على المحك، وهناك عدد كبير من المتغيرات (أسعار الفائدة، أسعار الصرف، التضخم، الإنتاج الصناعي، الصادرات، والبطالة) يجب أخذها في الاعتبار للوصول إلى حل لهذه المشكلة. كل هذه المتغيرات (وأكثر منها) تلعب دوراً في أي نقاش حول ما هو سعر الفائدة المناسب الذي يجب تطبيقه. وكمثال ثان، انظر إلى مسألة محاولة تفسير الاختلافات في أسعار المنازل، حيث يعتمد سعر المنزل على العديد من

المزايا (على سبيل المثال، عدد غرف النوم، عدد الحمامات، موقع المنزل، مساحة المنزل ... الخ). وكما هو موضح في المثال أعلاه، فإن هناك العديد من المتغيرات التي يجب تضمينها في النموذج بغرض توضيح لماذا تكون بعض المنازل أعلى سعراً من غيرها.

هذان المثالان يعبران عن حالة عامة لأن معظم المشكلات في علم الاقتصاد هي في نفس مستوى التعقيد وأكثر في ذلك. لسوء الحظ فإن الأداة الأساسية التي تعرفنا عليها حتى الآن، تحليل الارتباط البسيط، لا يمكنها التعامل مع هذا المستوى من التعقيد. ومن أجل التعامل مع حالات أكثر تعقيداً - مثل تلك التي تشتمل على أكثر من متغيرين - فإن الانحدار هو الأداة التي يجب استخدامها.

الانحدار كأفضل خط ملاءمة:

وكطريقة لفهم الانحدار، دعنا نبدأ بمتغيرين فقط (X,Y) حيث نشير لهذه الحالة بالانحدار البسيط. الانحدار المتعدد يتضمن العديد من المتغيرات، وسيتم مناقشته في الفصل السادس. البداية بالانحدار البسيط تعطي معنى ما دام بإمكاننا تطوير الإحساس البديهي من خلال الرسم البياني بطريقة مباشرة والعلاقة بين الانحدار والارتباط يمكن توضيحها بسهولة.

دعنا نرجع لأشكال الانتشار التي استخدمت سابقاً (على سبيل المثال، الشكل رقم 3-2) الذي يرسم أو يوضح الكثافة السكانية في مقابل الانحسار في مساحة الغابات، أو الشكل رقم (3-1) الذي يرسم أو يوضح مساحة المنزل في مقابل سعر المنزل). ناقشنا في الفصلين الثاني والثالث كيف أن دراسة أشكال الانتشار يمكن أن تكشف الكثير عن العلاقة بين X و Y. لذا وبصفة خاصة فإن رسم خط

مستقيم خلال النقاط على شكل الانتشار يقدم ملخصاً مناسباً للعلاقة بين X و Y. في تحليلات الانحدار، نقوم بتحليل هذه العلاقة بشكل منتظم.

وكبداية للموضوع، فإننا نفترض وجود علاقة خطية بين X و Y. وكمثال، يمكنك أخذ Y لتمثل المتغير سعر المنزل و X لتمثل المتغير مساحة المنزل من ملف البيانات HPRICE.xls. تذكر أن هذه البيانات تحتوي على الأسعار لـ 546 منزلاً في ويندسور بكندا وعلى مواصفات عديدة لكل منزل. وقد يكون من المناسب الافتراض بأن مساحة المنزل تؤثر في السعر الذي يباع به المنزل.

يمكننا التعبير عن العلاقة الخطية بين Y وX بطريقة رياضية كما يلي[1]:

$$Y = \alpha + \beta X$$

حيث α هي القاطع و β هي الميل. هذه المعادلة تعرف بخط الانحدار. وإذا كنا نعرف في الواقع ما هي قيم α وβ فإننا سوف نعرف ما هي طبيعة العلاقة بين X وY. وفي الواقع، فإننا لا نعلم ما هي طبيعة تلك العلاقة. إضافة لذلك، حتى إذا كان نموذجنا للانحدار الذي يوضح أن هناك علاقة خطية بين X و Y صحيحاً، فإننا لا نجد في الواقع العملي أن نقاط بياناتنا تقع بدقة على خط مستقيم. ولأسباب عدة منها على سبيل المثال "الخطأ في القياس /المقياس" فإن نقاط البيانات الفردية قد تقع بالقرب ولكن ليس تماماً على خط مستقيم.

على سبيل المثال، افترض أن سعر المنزل (Y) يعتمد على مساحته (X) وذلك على حسب التالي:

$$Y=34,000+7X$$

(حيث إن $\alpha = 34000$ و $\beta = 7$)

فإذا كانت X تساوي 5000 قدم مربع، فإن هذا النموذج يشير إلى أن سعر هذا المنزل يجب أن يساوي $Y = 34,000 + 7X\ (5,000) = \$69000$.

وبالطبع، فليس كل منزل بمساحة 5000 قدم مربع سوف يباع بسعر 69000 دولار أمريكي. ومما لا شك فيه فإنه في هذه الحالة سيفتقر نموذج الانحدار لبعض متغيراته المهمة (على سبيل المثال، عدد غرف النوم) والتي قد تؤثر في سعر المنزل. إضافة لذلك، فإن سعر بعض المنازل قد يكون أعلى مما يجب أن تكون عليه (على سبيل المثال، إذا تم شراؤهم من قِبل مشترين يملكون مالاً وفيراً وهم غير راشدين) وفي مقابل ذلك، فقد يتم بيع بعض المنازل بأقل من قيمتها الحقيقية (على سبيل المثال، إذا كان البائعون يرغبون في الانتقال لمدينة أخرى ويتوجب عليهم بيع منازلهم بسرعة). لكل هذه الأسباب، حتى لو كانت $Y=34,000+7X$ هي وصف دقيق لعلاقة خط مستقيم بين Y و X، فإنه لن تكون هناك حالة تقع فيها كل نقطة بيانات على نفس الخط.

مثالنا حول أسعار المنازل يوضح حقيقة مهمة حول تطبيقات نموذج الانحدار مفادها: أن نموذج الانحدار الخطي غالباً ما يكون وصفاً تقريبياً فقط للعلاقة الحقيقية. فالحقيقة قد تختلف بطرق عديدة من التقريب الذي يتضمنه نموذج الانحدار الخطي. في علم الاقتصاد، فإن أحد أكثر مصادر الخطأ شيوعاً هو بسبب المتغيرات غير الموجودة في النموذج، وغالباً بسبب عدم تمكننا من ملاحظتها. في مثالنا السابق، تعكس أسعار المنازل العديد من المتغيرات التي يمكننا جمع بياناتها بسهولة (على سبيل المثال، عدد غرف النوم، عدد الحمامات... الخ). إلا أنها تعتمد أيضاً على عناصر أخرى عديدة يصعب -إذا لم يكن من المستحيل- جمع بيانات عنها (على سبيل المثال، عدد الحفلات الصاخبة التي يقيمها الجيران، الدرجة التي يحفظ بها الملاك منازلهم بحالة جيدة، نوعية

الديكور الداخلي للمنزل ... الخ). استبعاد هذه المتغيرات من نموذج الانحدار يعني أن النموذج به خطأ، ونشير إلى جميع هذه الأخطاء بالحرف e.

يمكن الآن إعادة كتابة نموذج الانحدار على النحو التالي:

$$Y = \alpha + \beta X + e$$

في نموذج الانحدار، Y هي المتغير التابع، X هي المتغير التفسيري، α وβ هما معاملان. من الشائع الافتراض ضمنياً أن المتغير التفسيري "يتسبب" في تحديد قيمة Y ، وأن المعامل β يقيس تأثير X على Y.

في ضوء التعليقات الواردة في الفصل السابق بأن الارتباط لايعنى بالضرورة السببية، وقد ترغب في اختبار الافتراض بأن المتغير التفسيري يتسبب في تحديد قيمة المتغير التابع، هناك ثلاث إجابات يمكن صياغتها في هذا الشأن.

أولها: لاحظ أننا نتحدث عن نموذج الانحدار، هذا النموذج يحدد الكيفية التي تتفاعل بها المتغيرات المختلفة. على سبيل المثال، تشير نماذج استخدام الأرض أن الضغوط السكانية تتسبب في قيام المزارعين بتوسيع أراضيهم الزراعية من خلال قطع الغابات مما يتسبب في انحسار مساحة الغابات. مثل هذه النماذج لديها سببية (مضمنة في النموذج) والغرض من الانحدار الذي يشتمل على الانحسار في مساحة الغابات (Y)، الكثافة السكانية (X) هو قياس حجم أثر الضغوط السكانية فقط (على سبيل المثال، افتراض السببية قد يكون معقولاً لذلك فإننا نفترضه).

ثانيها: يمكننا التعامل مع الانحدار على أنه أسلوب لتعميم الارتباط وتفسير نتائج نموذج الانحدار على أساس أنها تعكس العلاقة بين المتغيرات. (بمعنى آخر، يمكننا إسقاط افتراض السببية إذا أردنا ذلك).

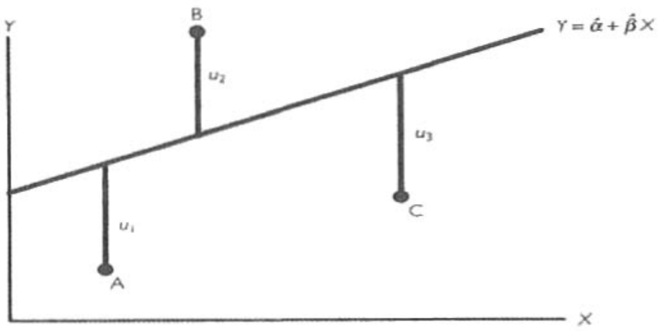
ثالثها: يمكننا الإقرار بأن الافتراض الضمني للسببية قد يكون مشكلة ومن ثم لابد من تطوير وسائل جديدة. هذا الموضوع سوف تتم مناقشته باختصار في الفصل الأخير من هذا الكتاب[2].

في ضوء معامل الخطأ، e، وحقيقة أننا لانعلم ماهى قيم α وβ، فإن المشكلة الأولى في تحليل الانحدار هى كيف يمكننا أن نحدد بالتقريب، أو نقدر، ماهى قيم α وβ. ويعتبر تسمية المقدرات α وβ كـ $\hat{\alpha}$ و $\hat{\beta}$ ممارسة قياسية أو عادية (بمعنى آخر، $\hat{\alpha}$ و $\hat{\beta}$ هما رقمان حقيقيان يحسبهما الحاسب الآلي، على سبيل المثال $\hat{\alpha} = 34.136$ و $\hat{\beta} = 6.599$، وهما تقديران للقيم الحقيقية غير المعروفة $\alpha = 34.000$ و $\beta = 7$. في الواقع العملي، فإن الطريقة التي نوجد بها التقديرات تتم من خلال رسم خط من خلال النقاط الواقعة على شكل الانتشار بحيث يمر بها بشكل أمثل.

لذلك لابد لنا من تعريف ماذا نعنى بـ "الخط الذي يربط بين النقاط بالشكل الأمثل". قبل أن نقوم بذلك، من المفيد أن نفرق بين الأخطاء (errors) والبواقي (residuals). الخطأ يعرف بأنه المسافة بين نقطة بيانات معينة وخط الانحدار الحقيقي. رياضياً، يمكننا إعادة ترتيب نموذج الانحدار لنكتب $e_i = Y_i - \alpha - \beta X_i$. وهو الخطأ للملاحظات i th. إلا أنه، إذا استبدلنا α وβ بتقديراتهما، $\hat{\alpha}$ و $\hat{\beta}$ فإننا سوف نحصل على خط مستقيم والذي بشكل عام يختلف قليلاً عن خط الانحدار الحقيقي. وتسمى الانحرافات عن خط الانحدار المقدر بالبواقي. وسوف نستخدم الحرف u_i عندما نشير للبواقي. أي أنه يمكن الحصول على البواقي من خلال المعادلة $u_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$. فإذا

وجدت أن التفريق بين الأخطاء والبواقي مربكاً، فيمكنك تجاهله في بقية هذا الكتاب وافتراض أن الأخطاء والبواقي هما شيء واحد. أما إذا كنت تخطط لمواصلة دراستك لعلم الاقتصاد القياسي، فإن هذا التفريق يصبح أساسياً.

وإذا رجعنا لبعض أساسيات الهندسة الخطية، فإنه يمكننا رسم خط مستقيم واحد (وواحد فقط) يربط بين أي نقطتين مختلفتين. لذلك، فإنه في حالة النقطتين، لا يوجد أي شك حول أفضل خط يمر بالنقطتين في شكل الانتشار. إلا أنه وبنفس المستوى لدينا العديد من النقاط – على سبيل المثال، مثالنا عن الانحسار في مساحة الغابات والكثافة السكانية يشتمل على 70 دولة مختلفة وعلى سبعين نقطة على شكل الانتشار – وهناك عدم وضوح حول "أفضل خط يربط بين النقاط". الشكل رقم (4-1) يحدد ثلاث نقاط بيانات (C,B,A) على شكل الانتشار. من الواضح أنه ليس هناك خط مستقيم يمر خلال هذه النقاط الثلاثة. فالخط الذي رسمته لا يمر بأي منها، بمعنى آخر، فإن كل نقطة تبعد قليلاً عنه.



شكل رقم (4-1) أفضل خط لثلاث نقاط

ولشرح ذلك بطريقة أخرى يمكن القول أن الخط الذي تم رسمه يشير إلى البواقي المشار إليها بـ u_1 وu_2 وu_3. والبواقي هي الفرق الرأسي بين نقطة البيانات والخط، والخط الذي يربط بين النقاط بالشكل الأمثل سيكون له بواق صغيرة.

الطريقة المعتادة لقياس حجم البواقي هي من خلال "مجموع مربعات البواقي" (Sum Squared Residuals) (SSR) كما يلي:

$$SSR = \sum_{i=1}^{N} u_i^2$$

حيث يتراوح عدد نقاط البيانات في العينة من i=1 ،......، إلى N. نحن نسعى للحصول على أفضل خط يربط بين النقاط والذي يقلل مجموع مربعات البواقي. ولهذا السبب، فإن التقديرات التي يتم إيجادها بهذه الطريقة تسمى تقديرات المربعات الصغرى (Least squares estimates) أو المربعات الصغرى الاعتيادية (Ordinary Least Squares) - OLS - للتفريق بينها وبين القياسات الأكثر تعقيداً والتي لن نقوم بمناقشتها إلا في الفصل الأخير من هذا الكتاب.

في الواقع، فإنه يمكن للبرامج مثل إكسل احتساب قيم $\hat{\alpha}$ و $\hat{\beta}$ وسوف يقلل ذلك مجموع مربعات البواقي. ويمكن اشتقاق المعادلة الصحيحة لكل من $\hat{\alpha}$ و $\hat{\beta}$ باستخدام التفاضل البسيط، ولكننا لن نقوم باشتقاقها هنا (انظر الملحق رقم (1-4) لمزيد من التفاصيل).

مثال: نموذج انحدار انحسار مساحة الغابات على الكثافة السكانية:

ارجع مرة أخرى لملف البيانات FOREST.XLs، والتي تشتمل على بيانات عن الكثافة السكانية والانحسار في مساحة الغابات لـ 70 دولة استوائية. وقد يكون من المناسب افتراض أن الكثافة السكانية تؤثر في الانحسار في مساحة الغابات وليس العكس. لذلك فإننا نختار الانحسار في مساحة الغابات كمتغير تابع (أي أن Y = معدل الانحسار في مساحة الغابات) والكثافة السكانية كمتغير تفسيري (أي أن X = الكثافة السكانية).

باستخدام برنـامج إكسل وأدواتـه (Tools/Data Analysis/Regression) نحصل على $\hat{\alpha} = 0.60$ و $\hat{\beta} = 0.00084$. وللتحدث بلغة أكثر علمية، لاحظ أننا عندما قدرنا نموذج الانحدار فإنه من الشائع القول "أننا نجري انحداراً لـ Y على X ".

لاحظ أيضا أنـه مـن الـسهل جداً حـساب هـذه الأرقـام فـي معظـم البـرامج الإحصائية. ومن المناسب أن نتحول بدلاً عن ذلك لقضية أكثر أهميـة وهـي: كيف يمكننا تفسير هذه الأرقام؟.

مثال: تكلفة إنتاج الكهرباء:

يحتـوى ملـف ELECTRIC.xls بيانـات عـن تكـاليف الإنتـاج (بملايـين الدولارات) لـ 123 شركة كهرباء في الولايات المتحدة الأمريكيـة فـي عـام 1970م. يتركـز الاهتمـام علـى فهـم العوامـل التـي تـؤثر فـي التكـاليف. وبذلك، Y = تكلفة الإنتاج وهى متغير تـابع. التكـاليف التـي تتحملها شـركة كهربـاء تعتمـد علـى العديـد مـن العوامـل، واحـد مـن أهـم هـذه العوامـل هـو بلاشك حجم الإنتاج (مقاساً بـآلاف الكيلووات ساعة من الكهربـاء المنتجـة) بواسـطة الـشركة. ونتوقـع أن تتحمـل الـشركات التـي تنـتج كهربـاء أكثـر تكاليف عالية (على سبيل المثال بسبب أنـه يتوجب عليها شراء وقود أكثر لإنتاج الكهرباء). وبذلك، X = الإنتاج كمتغير تفسيري مقبول. إذا أجرينـا الانحدار للتكاليف على الإنتاج، نحصل على $\hat{\alpha} = 2.19$ و $\hat{\beta} = 0.005$.

مثال: أثر الدعاية والإعلان فى المبيعات:

يشتمل الملف ADVERT.XLs على بيانات عن المبيعات السنوية ومصاريف الدعاية والإعلان (بملايين الدولارات) لـ 84 شركة في الولايات المتحدة الأمريكية. قد يكون المدير في الشركة مهتماً بمحاولة تحديد أثر الدعاية والإعلان في المبيعات. وهذا يتطلب إجراء انحدار باستخدام متغير تابع Y = المبيعات، ومتغير تفسيري X = مصاريف الدعاية والإعلان. وبذلك نحصل على القيم $\hat{\alpha} = 502.2$ و $\hat{\beta} = 0.218$، ويعد ذلك مؤشراً على وجود علاقة طردية بين الدعاية والإعلان والمبيعات.

تفسير تقديرات المربعات الصغرى الاعتيادية (OLS):

في المثال السابق حول العلاقة بين معدل الانحسار في مساحة الغابات والكثافة السكانية، حصلنا على تقديرات المربعات الصغرى الاعتيادية (OLS) لقاطع وميل خط الانحدار. والأن يبرز السؤال التالي: كيف يمكننا تفسير هذه التقديرات؟ القاطع في نموذج الانحدار α، عادة ما يكون لديه تفسير اقتصادي بسيط؛ لذلك لن نناقشه هنا. ومع ذلك، فإن β تعد مهمة جدا بنفس القدر. وهذا المعامل (β) هو ميل الخط المستقيم الذي يربط بين النقاط في شكل الانتشار.

في مثال الانحسار في مساحة الغابات/الكثافة السكانية، $\hat{\beta}$ كانت موجبة. وبالأخذ في الاعتبار النقاش حول كيفية تفسير الارتباطات في الفصل السابق نلاحظ أنه ما دامت $\hat{\beta}$ موجبة فإن X و Y في حالة ارتباط موجب. إلا أنه يمكننا أن نفسر $\hat{\beta}$ بشكل أكثر إذا احتسبنا التفاضل الرياضي لنموذج الانحدار وحصلنا على:

$$\frac{dY}{dX} = \beta$$

وحتى إذا كنت لاتعرف التفاضل، فإن من السهل تفسير دلالات هذه المعادلة. تقيس المشتقات حجم التغير في Y عندما تتغير X بمقدار صغير أو (حدي). حيث إن β يمكن تفسيرها كأثر حدي لـ X على Y وهو مقياس للكمية التي تتغير بها Y عندما تتغير X بوحدة واحدة[3].

يعتمد تعريف كلمة (وحدة) في الجملة السابقة بالتحديد على البيانات تحت الدراسة ويمكن توضيحها بشكل أفضل وتوضيحها خلال الأمثلة. وقبل القيام بذلك، يجب التأكيد على أن نماذج الانحدار تقيس اتجاهات البيانات (لاحظ استخدام الكلمة (اتجاه) في الشرح المتعلق بـ β أعلاه).

ليس من الضروري أن تتوافق كل مشاهدة من البيانات (على سبيل المثال، دولة أو منزل) مع النمط العام للبيانات. في الفصل الثاني سمينا هذه المشاهدات غير العادية بالقيم الشاذة (outliers) وقلنا إنه في بعض الحالات، يقدم اختبار القيم الشاذة معلومات كاملة. وفي حالة الانحدار، فإن القيم الشاذة هي الحالات التي تكون فيها البواقي ظاهرة لأنها كبيرة بشكل غير عادي. لذلك، فإن اختبار البواقي من الانحدار يُعد ممارسة شائعة. (في برنامج إكسل يمكنك اختبار البواقي بالنظر إلى الصندوق المعنون (Residuals) في قائمة اختيار الانحدار).

مثال: انحدار الانحسار في مساحة الغابات على الكثافة السكانية (تكملة ما ورد في الصفحات السابقة)

في مثال الانحسار في مساحة الغابات/ الكثافة السكانية حصلنا على $\hat{\beta} = 0.000842$. هذا يُعد مقياساً للقدر من انحسار مساحة الغابات الذي

يحدث عندما تتغير الكثافة السكانية بقدر قليل. وما دامت الكثافة السكانية يتم قياسها بعدد الأشخاص في كل 1000 هكتار، والانحسار في مساحة الغابات يتم قياسه كنسبة الفقد من إجمالي مساحة الغابات سنويا، فإن هذا الرقم يعنى أنه إذا أضفنا شخصاً واحداً في كل 1000 هكتار (بمعنى آخر، التغير بوحدة واحدة في المتغير التفسيري) فإن الانحسار في مساحة الغابات يميل للزيادة بنسبة 0.000842%.

وبصيغة أخرى، يمكننا إيضاح هذه المعلومات على النحو التالي: الكثافة السكانية تختلف كلياً في الدول مدار البحث: من أقل من 100 شخص لأكثر من 2500 شخص لكل 1000 هكتار. لذلك فإنه ليس من المستغرب أن تغيراً بمعدل شخص واحد لكل هكتار سوف يكون له أثر قليل في انحسار مساحة الغابات. يمكننا ضرب الكل بمقدار 100 ونقول إن زيادة الكثافة السكانية بمقدار 100 شخص لكل ألف هكتار سوف يميل لزيادة انحسار مساحة الغابات بمقدار 0.0842%. وحتى الرقم الأخير قد يبدو غير معنوي، إلا أنه يلاحظ أن زيادة معدلات انحسار مساحة الغابات السنوية بمقدار 0.0842% سنوياً سوف تكون نتيجتها خسارة الدولة 5% إضافية من غاباتها على مدى 50 عاماً. وعلى المدى الطويل، وعلى مساحة واسعة، وهي الاعتبارات المكانية والزمانية التي عادة ما يفكر بها الاقتصاديون البيئيون، فإن هذه الدرجة من انحسار مساحة الغابات قد تكون كبيرة.

مثال: إنتاج الكهرباء (تتمة ماورد في الصفحات السابقة)

في حالة انحدار التكاليف على الإنتاج حصلنا على $\hat{\beta} = 0.005$. بالأخذ في الاعتبار أن وحدات β هي قياس لأثر تغير المتغير التفسيري بوحدة واحدة

على المتغير التابع. وما دام الإنتاج يقاس بآلاف الكيلووات في الساعة (KWH)، فإن التغير بوحدة واحدة في المتغير التفسيري (Explonatory) يساوي ألف كيلووات/ ساعة. وحيث إن التكاليف تقاس بملايين الدولارات، فإن وحدات β هي β ملايين الدولارات. وبربط هذه الحقائق يمكننا القول بأن زيادة الانتاج بمقدار ألف كيلووات/ساعة تؤدى إلى زيادة التكاليف بمقدار 5000 دولار أمريكي (بمعنى آخر 0، 5000=1.000.000X0.005).

وبالطبع، يمكننا التعبير عن ذلك من ناحية الانخفاض بوحدة واحدة. أي أنه يمكننا القول أن خفض الإنتاج بمقدار 1000 كيلووات/ساعة سوف يؤدى لخفض التكاليف بمقدار 5000 دولار أمريكي.

مثال: أثر الدعاية والإعلان في المبيعات (تتمة ماورد في الصفحات السابقة)

تم قياس كل من الدعاية والإعلان والمبيعات بملايين الدولارات ووجدنا أن $\hat{\beta} = 0.218$. وبنفس أسلوب التبرير الوارد أعلاه، يمكننا القول بأن زيادة الدعاية والإعلان بمليون دولار تتسبب في زيادة المبيعات بمقدار 218.000 دولار (بمعنى آخر، 218.000=0.218X1.000.000). هذه النتيجة قد يبدو أنها تشير إلى أن الإنفاق على الدعاية والإعلان لايعد مجدياً ما دام إنفاق 1.000.000 دولار إضافية على الدعاية والإعلان يؤدي إلى زيادة قدرها 218.000 دولار في المبيعات.

هل يعني ذلك أن مسئول الشركة الذي يجري هذا الانحدار يتوجب عليه أن يقرر تخفيض مصاريف الدعاية والإعلان؟ من المحتمل، ولكنه ليس ضرورياً.

والسبب وراء عدم التأكد هذا يتعلق بقضية السببية، والسؤال عن كيفية تفسير نتائج الارتباط والانحدار (انظر الفصل الثالث أو بداية هذا الفصل).

أى أنه، إذا كان الانحدار حقيقة سببياً (بمعنى آخر، هي حالة أن يكون للدعاية والإعلان تأثير مباشر في المبيعات)، وبالتالي يمكننا القول أن مبلغ الـ 218.000 دولار يُعد مؤشراً على أثر أى تغير في الدعاية والإعلان. وإذا لم يكن ذلك سببياً، فإنه من المخاطرة استخدام نتيجة الانحدار لتقديم توصية إستراتيجية للشركة. بالطبع، قد تميل الشركات الكبيرة لإظهار نفسها بوصفها شركات كبيرة، إذ يستمتع مديروها برؤية شركاتهم يعلن عنها. وإذا كانت هذه القصة (التي ربما لاتكون مقنعة) صحيحة فسوف نتوقع أن نرى الشركات الكبيرة تعلن عن نفسها بشكل أكبر- وهو بالضبط ما توصل إليه انحدارنا. هذا التفسير قد يعنى أنه من الممكن ألا تؤثر الدعاية والإعلان في المبيعات بصورة مباشرة. العلاقة الإيجابية الواضحة بين الدعاية والإعلان والمبيعات من خلال تحليل الانحدار قد تكون متوافقة فقط مع سلوك مديري الشركات الكبيرة.

نظراً لصعوبة تحديد ما إذا كان من المناسب الافتراض بأن نموذج الانحدار يتضمن علاقة سببية يصعب فيها أن يؤثر أحد المتغيرات بشكل مباشر في المتغير الأخر، وكذلك يصعب تقديم أحكام عامة عن الموضوع، قد يكون من الأفضل صياغة قاعدة عامة ونظرية اقتصادية للاسترشاد بها في الشرح والتفسير.

تمرين رقم (4-1):

يحتوي ملف برنامج إكسل FOREST.XLS على بيانات Y= معدل الانحسار في مساحة الغابات، X= الكثافة السكانية، W= التغير في مساحة الأراضي المخصصة لزراعة المحاصيل Z= التغير في مساحة الأراضي المخصصة للمراعي.

(أ) قم بإجراء انحدار لـY على X واشرح النتائج.

(ب) قم بإجراء انحدار لـY على W وآخر لـ Y على Z واشرح النتائج.

(جـ) أوجد متغيراً جديداً، V، من خلال قسمة X على 100. ما هي طبيعة الوحدات التي تم قياس V على أساسها؟

(د) قم بإجراء انحدار لـY على V. قارن النتائج التي حصلت عليها بنتائج الانحدار في الفقرة (أ) أعلاه. كيف تفسر تقديرك للمعامل β؟ كيف تختلف $\hat{\alpha}$ بين الفقرتين (أ) و (د)؟

(هـ) قم بتعديل المتغيرات التابعة والمستقلة (بمعنى آخر، قسمة كل منهما على مقدار ثابت) ولاحظ الأثر الذي يتركه ذلك على تقديرك للمعامل.

القيم المقدرة و R^2: قياس ملاءمة نموذج الانحدار:

من النقاش السابق عرفنا كيف نحسب ونفسر معاملات الانحدار، $\hat{\alpha}$ و $\hat{\beta}$. إضافة لذلك، فقد أوضحنا أن الانحدار يحدد "أفضل خط ملائم". أي أنه يقلل قيمة SSR. ومع ذلك، فإنه من الممكن أن يكون الخط "الأفضل" ملاءمة غير ملائم تماماً. ومن المناسب أن يكون لدينا قياس للملاءمة. (أو قياس للكيفية التي يكون عليها أفضل خط ملائم جيداً). وتُعد R^2 أكثر مقياس للملاءمة شيوعاً. ويتعلق R^2 بشكل لصيق بالارتباط بين Y و X. وهو عبارة عن مربع قيمة معامل الارتباط في نموذج الانحدار البسيط، وهو الرابط الإحصائي بين الانحدار والارتباط. ومع ذلك، فإن النقاش السابق يجب أن يجعل الروابط غير العادية بين

الارتباط والانحدار واضحة. وكلاهما يهتمان بتحديد درجة العلاقة بين مختلف المتغيرات ويمكن تفسيرهما من خلال الخطوط الملائمة في شكل الانتشار.

لاشتقاق وتوضيح R^2 ، لا بد من تقديم خلفية مبسطة، نبدأها بتوضيح مصطلح القيمة الملائمة. تذكر أن الانحدار يلائم الخط المستقيم في شكل الانتشار، إلا أنه لا يمر بدقة من خلال كل نقطة على الانتشار (بمعنى آخر، هناك خطأ قد حدث). في حالة مثالنا عن الانحسار في مساحة الغابات/الكثافة السكانية، نجد أن الدول منفردة لا تقع على خط الانحدار المقابل لقيمة X_i لتلك المشاهدة المحددة (على سبيل المثال، المنزل، الدولة). بمعنى آخر، إذا رسمت خطاً عمودياً مستقيماً خلال نقطة معينة في شكل الانتشار، فإن تقاطع هذا الخط العمودي وخط الانحدار هو القيمة الملائمة التي تقابل النقطة التي اخترتها.

وبدلاً عن ذلك، يمكننا التفكير في فكرة مقدار الملاءمة في شكل معادلة نموذج الانحدار

$$Y_i = \alpha + \beta X_i + e_i$$

لاحظ أن إضافة علامة i (على سبيل المثال Y_i) تشير إلى أننا نشير إلى مشاهدة معينة (على سبيل المثال الدولة ith أو المنزل ith). إذا تجاهلنا الخطأ، يمكننا القول بأن توقع (تنبؤ) النموذج لـ Y_i يجب أن يكون مساوياً لـ $\alpha + \beta X_i$.

وإذا قمنا باحلال β,α بتقديرات المربعات الصغرى الاعتيادية (OLS) $\hat{\alpha}$ و $\hat{\beta}$، نحصل على ما يسمى "قيمة ملائمة" أو "قيمة تم توقعها" لـ Y_i

$$\hat{Y}_i = \hat{a} + \hat{\beta} X_i$$

لاحظ أننا نستخدم قيمة المتغير التفسيري وتقديرات المربعات الصغرى الاعتيادية لتوقع قيمة المتغير التابع. ومن خلال النظر لقيم (Y_i) الحقيقية مقابل $\hat{Y}_i$ الملائمة يمكننا الحصول على انطباع مبدئي حول "مدى دقة الملاءمة" لنموذج الانحدار.

لاحظ أن العديد من برامج الحاسب الآلي تتيح لك طباعة القيم الحقيقية والملائمة لكل مشاهدة. واختبار هذه القيم لايعطيك قياساً مبدئياً فقط للكيفية التي يكون بها النموذج ملائماً، وإنما يتيح لك اختبار المشاهدات المفردة لتحديد المشاهدات أو النقاط الأكثر قربا لخط الانحدار والمشاهدات التي ليست قريبة منه. ولأن خط الانحدار يشتمل على الأنماط العامة أو التوجهات في مجموعة بياناتك، فإنه يمكنك معرفة أى من المشاهدات تتلاءم مع النمط العام والتي لا تتلاءم مع ذلك النمط.

تمرين رقم (2-4):

باستخدام البيانات المتوافرة في الملف FOREST.xls (انظر تمرين رقم 1-4) قم بإجراء انحدار لـ Y على X مستخدماً برنامج إكسل بالضغط على المربع المعنون (Line Fit Plot) في قائمة الانحدار. قارن بيانياً ورقمياً القيم الحقيقية والملائمة (بمعنى آخر، انظر للأعمدة المعنونة Residual output والرسم المصاحب لها).

لقد عرفنا سابقا البواقي التي تنتج من عملية ملاءمة خطنا لأفضل قيم ملائمة. وهناك طريقة أخرى للتعبير عن هذا الباقي تتمثل في شكل الفرق بين قيم Y الحقيقية والملائمة أى أن:

$$u_i = Y_i - \hat{Y}_i$$

إن برامج الحاسب مثل إكسل يمكنها أيضا رسم أو عرض البواقي من نموذج الانحدار. وهذه البواقي يمكن اختبارها في المقابل لإعطاء انطباع مبدئي لدرجة ملائمة لنموذج الانحدار. ونؤكد بأن البواقي الكبيرة بشكل غير عادى تكون بعيدة عن خط الانحدار، وأن القيم الشاذة في بعض الأحيان تكون محل اهتمام.

تمرين رقم (4-3):

(أ) باستخدام البيانات في الملف FOREST.xls (انظر تمرين رقم 4-1) قم بإجراء انحدار لـ Y على X مستخدماً برنامج إكسل بالنقر على المربعات المعنونة "Residuals" و "Residual Plots" في قائمة الانحدار.

كيف يمكنك تفسير البواقى؟ هل هناك أى نقاط بعيدة عن خط الانحدار؟

(ب) كرر السؤال (أ) بالنسبة للمتغيرات، W و Z في مجموعة البيانات.

ولتوضيح نوع المعلومات التى يمكن أن يقدمها لنا تحليل البواقي، انظر للنتيجة التى توصلت إليها عن طريق الحاسب الآلي من الفقرة (أ) في التمرين رقم (4-3).

في ناتج البواقي نجد أن المشاهدة رقم 39 لها قيمة ملائمة قدرها 2.93 وباقي بمقدار -1.63. وبإضافة هذين الرقمين لبعضهما (أو من خلال النظر إلى البيانات الأصلية)، يمكنك معرفة أن معدل الانحسار في مساحة الغابات الحقيقي لهذه الدولة هو 1.3. ماذا تعني كل هذه الأرقام؟ لاحظ أن نموذج الانحدار يتوقع

قيمة أعلى للانحسار في مساحة الغابات تصل إلى (2.93) أكثر مما هو حاصل فعلياً في هذه الدولة (1.3). وهذا يعني أن هذه الدولة عملت بشكل جيد لحماية غاباتها مقارنة بما يشير إليه نموذج الانحدار، وتبعاً لذلك، فهى تقوم بمجهود فعال للمحافظة على الغابات مقارنة بالدول الأخرى، وأن هذه الدولة المتفردة قد تقدم دروساً مفيدة يمكن تطبيقها على بقية الدول.

تعد المفاهيم المتعلقة بموضوعي البواقي والقيمة الملائمة مهمة في تطوير فهم غير منهجي حول كيفية أن يكون نموذج الانحدار ملائماً بشكل جيد. ومع ذلك، فإننا مازلنا نفتقر لمقياس رقمي منهجي للملاءمة. في هذه المرحلة، يمكننا الآن اشتقاق وتطوير المقياس R^2، آخذين في الاعتبار أن التباين هو مقياس الانتشار أو الاختلاف في البيانات. في هذا السياق نقوم بتعريف مفهوم يرتبط بذلك ارتباطاً وثيقاً، ألا وهو إجمالي مجموع المربعات أو Total Sum of Squares (TSS).

$$TSS = \sum (Y_i - \bar{Y})^2,$$

لاحظ أن معادلة التباين لـ Y هى ⟵ TSS/N-1 (انظر الفصل الثاني).

وبشكل عام فإن المقدار N-1 سوف يتم إلغاؤه في معادلتنا الأخيرة لـ R^2 وتبعاً لذلك سوف نتجاهله. لذلك فكر في TSS كمقياس للتغير في Y. يهدف نموذج الانحدار لتوضيح التغير في Y من خلال المتغير التفسيري X. ويمكن الإشارة إلى أن التغير الكلي في Y يمكن تقسيمه إلى جزأين كما يلي:

TSS=RSS+SSR

حيث RSS هي مجموع مربعات الانحدار (Regression Sum of Squares)، وهو مقياس للقدرة التفسيرية لنموذج الانحدار [4] RSS، ويمكن الحصول على RSS من خلال الصيغة التالية:

$$RSS = \sum(\hat{Y}_i - \bar{Y})^2$$

وبالأخذ في الاعتبار أن SSR هي مجموع البواقي المربعة وأن نموذج الانحدار الموافق (الملائم) بشكل جيد سوف يجعل من قيمة SSR صغيرة جداً، فيمكننا الجمع بين المعادلات أعلاه للحصول على مقياس الملاءمة التالي:

$$R^2 = 1 - \frac{SSR}{TSS}$$

أو بنفس القدر:

$$R^2 = \frac{RSS}{TSS}$$

بديهياً، تقيس R^2 النسبة من إجمالي تباين Y التى يمكن تفسيرها من خلال X. تذكر أن SSR,RSS,TSS جميعها مجاميع لأرقام مربعة و تبعاً لذلك، جميعها غير سالبة.

وهذا يعني أن $TSS \geq RSS$ و $TSS \geq SSR$. باستخدام هذه الحقائق يمكن رؤية أن $0 \leq R^2 \leq 1$.

ويمكن الحصول على استنتاج أكثر حول مقياس الملاءمة هذا بالأخذ في الاعتبار أن القيم الصغيرة لـ SSR تشير إلى أن نموذج الانحدار متوافق بشكل جيد. خط الانحدار الذي يتوافق كلياً مع كل نقاط البيانات بشكل انتشار ليس فيه أخطاء، لذلك فإن SSR=0 و $R^2 = 1$. وبالنظر إلى المعادلة أعلاه، يمكنك أن ترى أن قيم R^2 القريبة من 1 تعني ملاءمة جيدة وأن $R^2 = 1$ تعني ملاءمة تامة. إجمالاً، يمكن القول أن القيم العالية لـ R^2 تعني ملاءمة جيدة وأن القيم المنخفضة تعني ملاءمة غير جيدة. RSS تقدم تفسيراً بديهياً آخر لمدى ملاءمة

الانحدار وتقيس RSS كيف يمكن أن تفسر المتغيرات المستقلة "التغيرات" في قيمة Y. فإذا كانت RSS قريبة لـ TSS، فإن المتغيرات التفسيرية تكون مسئولة تقريباً عن كل التغير في قيمة Y وتكون الملاءمة جيدة. وبالنظر للمعادلة السابقة فسوف تلاحظ أن قيمة R^2 تقترب من الواحد صحيح في هذه الحالة.

مثال: تكلفة إنتاج الكهرباء (تكملة ما ورد في الصفحات السابقة)

في انحدار Y = تكلفة الإنتاج على X = الإنتاج لـ 123 شركة كهرباء، $R^2 = 0.92$. وهذا رقم عال جداً وقريب من الواحد صحيح، مما يشير إلى أن ملاءمة خط الانحدار جيدة جداً. بمعنى آخر، فإن 92% من الاختلاف في التكاليف في كل الشركات يمكن تفسيره من خلال الاختلاف في الإنتاج. لاحظ أنه إذا قمت بحساب الارتباط بين الإنتاج والتكلفة فإنك سوف تحصل على $r_{xy} = 0.96$. قيمة هذا الارتباط المربعة تساوي تماماً R^2 (بمعنى آخر، $0.92 = 0.96^2$).

هذا المثال يلقي الضوء على العلاقة القوية بين الارتباط والانحدار. لاحظ أن قيمة R^2 الناتجة من انحدار Y على X تساوي تماماً مربع قيمة الارتباط بين Y و X. الانحدار هو حقيقة مجرد امتداد للارتباط. حتى الآن، يقدم الانحدار أيضا تعبيراً واضحاً للأثر الحدي (β)، الذي غالباً ما يكون مهماً بالنسبة للتحليلات المتعلقة بوضع السياسات الإنتاجية.

مثـال: أثـر الدعايـة والإعـلان فـي المبيعـات (تكملـة مـا ورد فـي الـصفحات السابقة)

قيمة R^2 الناتجة عن انحدار المبيعات على تكاليف الدعاية والإعلان باستخدام البيانـات فـي الملف ADVER.xls هـي 0.09. هذا الرقم الصغير نـسبياً يشير إلـى أن الفروقـات فـي تكـاليف الدعايـة والإعلان فـي كل الشركات تمثل نـسبة صغيرة فقط من الاختلاف فـي المبيعـات. هذه النتيجـة التـى توصلنا إليها ربمـا تكـون مناسبة، لأنك قد تتوقع أن تلعب عوامل أخـرى غيـر الدعايـة والإعـلان (علـى سبيل المثـال، نوعيـة الإنتـاج، التـسعير ... الـخ) دوراً مهمـاً فـي توضيح وضع المبيعات في الشركة.

تمرين رقم (4-4):

(أ) باستخدام البيانات في الملف FOREST.xls (انظر التمرين رقم 4-1)، قم بإجراء انحدار لـ Y على X باستخدام برنامج إكسل. ماهى قيمة R^2 ؟

(ب) احسب الارتباط بين Y و X.

(ج) ناقش العلاقة بين النتائج في كل من (أ) و (ب) أعلاه.

(د) أعد تطبيق السؤال (أ) أعلاه لمختلف الانحدارات التى تتضمن Z,Y,X,W فـي مجموعـة البيانـات. علـق علـى وضـع الملاءمـة (Fit) لكـل هـذه الانحدارات.

عدم الخطية في الانحدار:

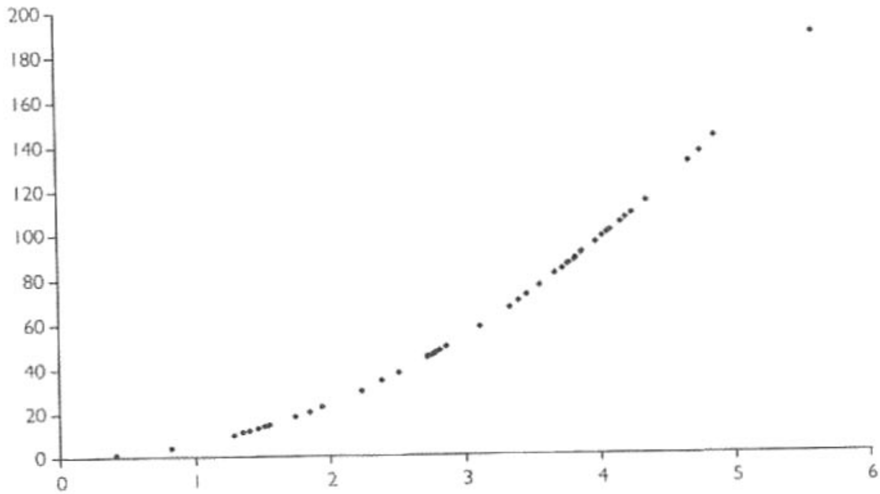
حتى الآن، قمنا باستخدام نموذج الانحدار الخطي وملاءمة الخط المستقيم خلال شكل الانتشار. إلا أن ذلك قد لايكون مناسباً دائماً. انظر إلى الشكل رقم (2-4) إنه يشير إلى أن العلاقة بين Y و X ليست خطية. فإذا كنا نريد ملاءمة الخط المستقيم من خلال البيانات، فإن ذلك قد يعطي تمثيلاً مضللاً للعلاقة بين Y و X. في الحقيقة أننا قمنا باستخراج هذه البيانات بطريقة غير حقيقية من خلال افتراض أن العلاقة بين Y و X هي على الشكل:

$$Y_i = 6X_i^2,$$

حيث تكون العلاقة الحقيقية تربيعية (من الدرجة الثانية). وتشير النظرة الفاحصة لشكل الانتشار غالباً إلى ما إذا كانت ملاءمة الخط المستقيم مناسبة أم لا.

ماذا تفعل إذا كانت العلاقة عبارة عن معادلة رياضية من الدرجة الثانية أكثر من أنها معادلة خطية؟ الإجابة عن ذلك سهلة وتتمثل في إجرائك انحدار لـ Y على X^2 بدلاً عن إجراء انحدار لـ Y على X.

بالطبع، فإن العلاقة التى تم توضيحها من خلال شكل الانتشار قد لاتكون خطية ولا معادلة رياضية من الدرجة الثانية. قد يبدو أن Y لها علاقة بـ $\ln(X)$ أو $1/X$ أو X^3 أو أى تحويل آخر لقيم X. ومع ذلك، فإن نفس القاعدة العامة تنطبق: قم بتحويل المتغير X بحسب ما يتناسب وبعد ذلك قم بإجراء انحدار لـ Y على المتغير الذي تم تحويله. يمكنك أيضا تحويل Y إذا كان ذلك مناسباً.



شكل رقم (2-4) العلاقة التربيعية بين Y,X

يُعد التحويل اللوغريثمي لكل من المتغيرات التابعة والتفسيرية تحويلاً شائعاً جداً. حتى إذا لم تكن على علم باللوغاريثمات، فإنه من السهولة العمل عليها في برامج الحاسب الآلي للاقتصاد القياسي بما فيها إكسل[5]. غالباً ما يعمل الاقتصاديون باللوغاريثمات الطبيعية، التى رمزها $\ln$. في هذا الكتاب، سوف نستخدم دائماً اللوغاريثمات الطبيعية وسوف نشير إليها اختصارا بـ (Logs). من الشائع أن نقول: "أخذنا Log المتغير X" أو أن نقول "أننا عملنا باستخدام Log X " الذي تعريفه الرياضي هو $\ln(X)$.[6]

لماذا من الشائع استخدام $\ln(Y)$ كمتغير تابع و $\ln(X)$ كمتغير تفسيري؟

أولاً: لأنه يمكننا بهذين التعبيرين تفسير النتائج التي توصلنا إليها بطريقة سهلة جداً.

ثانياً: البيانات التى تم تحويلها بهذه الطريقة تبدو كافية،غالباً، للإيفاء بافتراض خطية نموذج الانحدار.

لفهم النقطة الأولى بوضوح، فإننا نحتاج إلى خلفية في حساب التفاضل والتكامل، وهو أمر خارج نطاق هذا الكتاب. ولحسن الحظ، فإنه يمكن التعبير عن فهمنا للأمر دون الخوض في تفاصيله الرياضية، خذ على سبيل المثال معادلة الانحدار التالية:

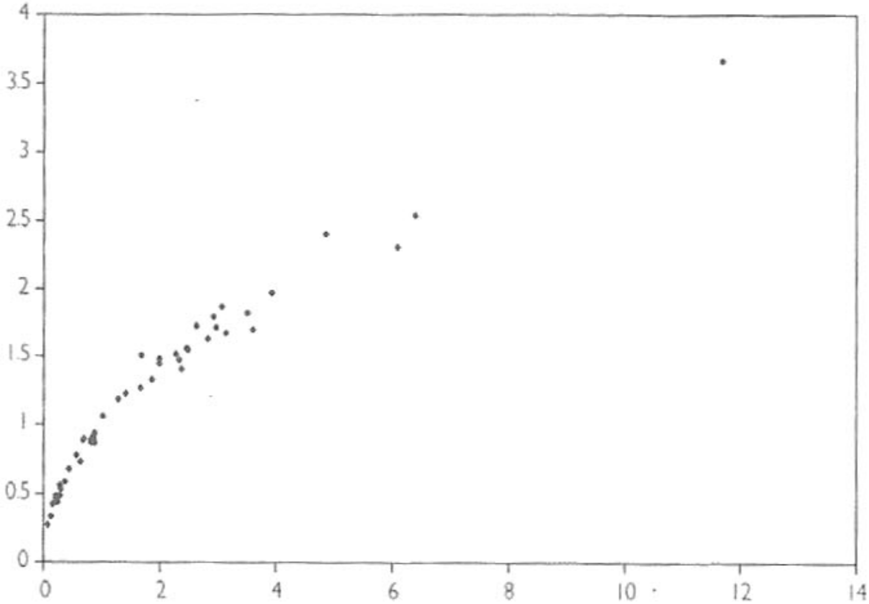
$$\ln(Y) = \alpha + \beta \ln(X) + e,$$

حيث β هي المرونة، آخذين في الاعتبار أننا قلنا أنه في الانحدار الأساسي بدون لوغاريثمات طبيعية أن "Y تميل للتغير بوحدات β لكل تغير بوحدة واحدة في X". في الانحدار الذي يشتمل على كل من المتغيرات التابعة والتفسيرية المعبر عنها بلوغاريثمات طبيعية، يمكننا الآن القول أن Y تميل للتغير بنسبة مئوية من β لكل تغير بنسبة مئوية واحدة من X. أي أنه، بدلاً من القلق حول وحدات القياس، فإن نتائج الانحدار باستخدام المتغيرات اللوغاريثمية الطبيعية غالباً ما يتم التعبير عنها كمرونات.

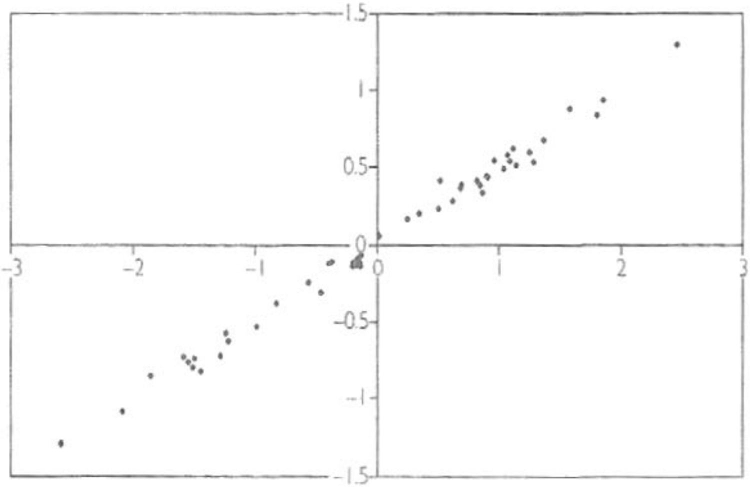
اللوغاريثمات الطبيعية مناسبة أيضاً لأسباب أخرى. على سبيل المثال، كما تمت المناقشة في الفصل الثاني، عندما يكون لدينا بيانات سلاسل زمنية، فإن التغير النسبي في المتغير هو بالتقريب $[In(Y_t) - In(Y_{t-1})]$ 100x. هذا التحويل سوف يكون مفيداً في الفصول اللاحقة من هذا الكتاب.

التبرير الثاني للتحويل للوغاريثمات هو تبرير عملي خالص: في حالات كثيرة، إذا أخذت اللوغاريثم الطبيعي لكل من المتغيرات التابعة والتفسيرية ورسمت شكل الانتشار (إحداثيات XY) فإن العلاقة الناتجة سوف تبدو خطية. هذا ما تم توضيحه في الشكل رقم (4-3) والشكل رقم (4-4). الشكل رقم (4-3) هو عبارة عن شكل الانتشار لمجموعتي بيانات سلاسل زمنية، Y , X علماً أنه لم يتم تحويل أي من Y X، بأي طريقة.

الشكل رقم (4-4) هو عبارة عن شكل الانتشار لكل من $\ln(Y), \ln(X)$. لاحظ أن النقاط في الشكل الأول لا تبدو أنها تقع على الخط المستقيم. وبدلاً عن ذلك فإن العلاقة هي علاقة ذات نمط ميل متدرج للقيم الصغيرة من X، وأنها تقترب تدريجياً من الخطية كلما ارتفعت قيم X. هذا نمط نموذجي للبيانات التي يجب تحويلها إلى لوغاريثمات طبيعية. الشكل رقم (4-4) يوضح أنه بمجرد وضع البيانات في شكل لوغاريثمات طبيعية، فإن شكل الانتشار يشير إلى نمط خطي. انحدار المربعات الصغرى الاعتيادية (OLS) سوف يتلاءم مع خط مستقيم بدقة عالية في الشكل رقم (4-4). وبالرغم من ذلك، فإن ملاءمة خط مستقيم دقيق من خلال الشكل رقم (4-3) هو أمر صعب للغاية (وربما من الأفضل عدم القيام به).



شكل رقم (4-3) X و Y تحتاج لأن توضع في شكل لوغريثمى



شكل رقم (4-4) $\ln(X)$ **في مقابل** $\ln(Y)$

على أي أساس يجب عليك تحويل بياناتك إلى لوغاريثمات طبيعية (أو أي تحويل آخر)؟ ليس هناك قاعدة عامة يمكن إتباعها. أحد الطرق المتبعة هو مقارنة شكل الانتشار لبياناتك الأصلية ومقارنتها مع أشكال انتشار أخرى بعد تحويلها بطرق مختلفة. على سبيل المثال، ابدأ بالنظر إلى انتشار X في مقابل Y، سيبدو خطياً بشكل عام. وإذا كان كذلك، قم بإجراء انحدار Y على X. إذا كان الانتشار لا يبدو خطياً، فربما يوضح نمطاً آخر تعرفه (على سبيل المثال، شكل المعادلة الرياضية من الدرجة الثانية في الشكل رقم 4-2 أو الشكل اللوغريثمي للشكل رقم 4-3). إذا كان الأمر كذلك، ارسم شكل انتشار لمتغيرات مناسبة تم تحويلها (على سبيل المثال، $\ln(Y)$ في مقابل $\ln(X)$) وانظر إذا كانت تبدو خطية. هذه الطريقة قد تكون فعالة في الانحدار البسيط الذي يشتمل على متغير تفسيري واحد. في الفصل السادس، سوف نناقش حالات تتضمن متغيرات تفسيرية عديدة، وفي هذه الحالات قد يكون اختبار شكل الانتشار معقداً بشكل كبير ما دام هناك العديد من أشكال الانتشار يمكن رسمها.

تمرين رقم (4-5):

باستخدام البيانات في الملف FOREST.xls قم برسم أشكال انتشار تشتمل على Z,W,Y,X (انظر تمرين رقم 4-1 من أجل الحصول على تعريف لهذه المتغيرات)

هل يبدو أن هناك علاقة غير خطية بين أي اثنين من هذه المتغيرات؟

كرر التمرين باستخدام البيانات في مثال الدعاية والإعلان على الملف (ADVERT.xls)

تمرين رقم (4-6):

تحتوى البيانات في الملف EX46.xls على متغيرين هما X,Y.

(أ) ارسم شكل انتشار لهذين المتغيرين. هل تبدو العلاقة بين X,Y خطية.

(ب) احسب الجذر التربيعي للمتغير X. لاحظ أن الدالة (Function) المستخدمة في برنامج إكسل للجذر التربيعي هى SQRT.

(ج) ارسم شكل انتشار للجذر التربيعي لـ X في مقابل Y. هل تبدو هذه العلاقة خطية؟

تمرين رقم (4-7):

استخدم بيانات المثال المتعلق بتكاليف إنتاج الكهرباء (ELECTRIC.xls)، حيث Y= تكلفة الإنتاج و X = حجم الإنتاج

(أ) قم بإجراء انحدار لـ Y على X.

(ب) احسب تحويلات اللوغريثم الطبيعي لكلا المتغيرين.

(ج) قم بإجراء انحدار لـ $\ln(Y)$ على $\ln(X)$ واشرح النتائج التى توصلت إليها.

ملخص الفصل

1- الانحدار البسيط يحدد كمياً / رقمياً أثر المتغير التفسيري X، على المتغير التابع Y. لذلك فهو يقيس العلاقة بين متغيرين.

2- العلاقة بين Y و X يفترض أن تأخذ الشكل $Y = \alpha + \beta X$، حيث α هى القاطع و β هى ميل الخط المستقيم. وهذا يسمي خط الانحدار.

3- خط الانحدار هو أحسن خط ملاءمة (Fitting) خلال شكل الانتشار.

4- ليس هناك خط ملاءمة بشكل كامل خلال كل النقاط في الرسم البياني XY. المسافة بين كل نقطة والخط تسمى الباقي.

5- تقدير المربعات الصغرى الاعتيادية (OLS) هو القياس الذي يخفض قيمة مجموع البواقي المربعة إلى حدها الأدنى.

6- تقدم OLS مقدرات لكل من α وβ وتعرف بـ $\hat{\alpha}$ و $\hat{\beta}$.

7- من الملائم تفسير معاملات الانحدار على أنها آثار حدية (بمعنى آخر، كقياسات للأثر على Y نتيجة لتغير صغير في X).

8- R^2 هي قياس لمدى الملاءمة الجيدة لخط الانحدار خلال الرسم البياني XY.

9- مقدرات OLS و R^2 يتم قياسهما أو احتسابهما في برامج الحاسب الآلي مثل إكسل.

10- خطوط الانحدار لا تكون دائماً خطية. لرسم خط انحدار غير خطي، قم بإبدال Y أو X أو كليهما في نموذج الانحدار بتحويل غير خطي مناسب (على سبيل المثال، $\ln(Y)$ أو X^2).

ملحق رقم (4-1): التفاصيل الرياضية:

يُعرف مقياس OLS بأنه أحسن خط ملاءمة يمر خلال النقاط في شكل الانتشار. رياضياً، نحن نهتم باختيار $\hat{\alpha}$ و $\hat{\beta}$ حتى نتمكن من تصغير (Minimize) مجموع البواقي المربعة.

يمكن كتابة SSR (Sum of Squared Residual) كما يلي:

$$SSR = \sum_{i=1}^{N}(Y_i - \hat{\alpha} - \hat{\beta} X_i)^2.$$

تمرين اختياري:

احصل على الاشتقاق الأول والثاني بدلالة $\hat{\alpha}$ و $\hat{\beta}$ لمعادلة SSR أعلاه. واستخدمهما للحصول على قيم $\hat{\alpha}$ و $\hat{\beta}$ التى تخفض قيمة SSR. اثبت أن الحل الذى توصلت إليه لايخفض بالضرورة (إنما قد يرفع) قيمة SSR.

إذا قمت بإجراء التمرين السابق بطريقة صحيحة، فإنك لابد أن تكون قد حصلت على ما يلي:

$$\hat{\beta} = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{N}(X_i - \bar{X})^2}$$

و

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X},$$

حيث $\bar{Y}$ و $\bar{X}$ هما المتوسط الحسابي لكل من Y و X (انظر الفصل الثاني) وهى مقدرات المربعات الصغرى الاعتيادية (OLS) لكل من α وβ. لاحظ أن هناك العديد من الطرق المماثلة لكتابة معادلة $\hat{\beta}$. إذا اطلعت على مراجع أساسية أخرى حول الموضوع سوف تجد معادلات مختلفة لقياس OLS.

هذه المعادلات يمكن استخدامها لإيضاح تبعات أخذ الفروقات من المتوسطات الحسابية.

وفي سياق الشرح والتوضيح، لاحظ أننا افترضنا أعلاه أن المتغيرات التابعة والمستقلة X و Y، تقوم على البيانات غير المعالجة (الأصلية). إلا أن الباحثين لا يتعاملون في بعض الحالات مع قيم X و Y المجردة، وإنما مع قيم X و Y مطروحاً منها قيم متوسطاتها الحسابية كما يلي:

$$y_i = Y_i - \bar{Y}$$

و

$$x_i = X_i - \bar{X}$$

والآن يمكنك دراسة استخدام OLS لقياس الانحدار على النحو التالي:

y=a+bX+e,

حيث تم استخدام a و b للتفريق بينهما وبين المعاملات α وβ في الانحدار الذي يتضمن Y و X.

وفي المقابل، فإن العلاقة بين تقديرات OLS من الانحدار الأصلي وتلك ذات الانحرافات المأخوذة من الأوساط الحسابية هى علاقة بسيطة. تقدير OLS لـ b

هو دائماً يساوي تماماً $\hat{\beta}$ وقياس OLS لـ a هو دائماً يساوي صفراً. بمعنى آخر، يؤدي أخذ الانحرافات من الأوساط الحسابية إلى تبسيط نموذج الانحدار من خلال التخلص من القاطع (بمعنى آخر، ليس هناك معنى لتضمين القاطع ما دام معامله يساوي دائماً صفراً). هذا التبسيط ليس له أي أثر على معامل الميل في نموذج الانحدار، وهو لا يتغير من خلال أخذ الانحرافات من الأوساط الحسابية ويظل لديه نفس المعنى كأثر حدي.

ولا يعتبر إثبات الجمل في الفقرة السابقة صعباً جداً، فإذا كنت ملماً بالرياضيات، فقد ترغب في القيام بذلك. وكخطوة أولية لبداية العمل تذكر أن الأوساط الحسابية لكل من y وx هي صفر.

في الفصل السادس، سوف نناقش حالات ذات متغيرات تفسيرية متعددة. وفي هذه الحالة، إذا أخذت الانحرافات من الأوساط الحسابية من المتغير التابع وكل المتغيرات التفسيرية، فإنك سوف تحصل على نفس النتيجة. أي أن القاطع سوف يختفي من الانحدار، إلا أن كل قياسات المعامل سوف تبقى كما هي.

ملاحظات ختامية:

1- لاحظ أننا تجاهلنا علامة الضرب في الكثير من المواقع بغية التبسيط.

وعلى سبيل المثال بدلاً من القول $Y = \alpha + \beta x X$، فإننا فقط نقول $Y = \alpha + \beta X$.

2- بعض الكتب الإحصائية تحدد خطاً فاصلاً بين الارتباط والانحدار. وتشير هذه الكتب إلى أن الارتباط يجب أن يتم التعبير عنه كمقياس للعلاقة بين متغيرين، وليس للسببية. وفي المقابل، الانحدار يجب أن يقوم على السببية في شكل عبارات مثل: "النظرية الاقتصادية تخبرنا أن X تتسبب في Y". بالطبع فإن هذا التقسيم يبسط تفسير مدلولات النتائج القياسية.

وإجمالاً، يمكن القول بأن ذلك يسهل عملية التفكير في المتغير التابع، الذي تم وضعه على جانب واحد من معادلة الانحدار، باعتباره ناتجاً من قبل المتغيرات التفسيرية على الجانب الأخر من المعادلة. ومع ذلك، يمكن القول بأن هذا التقسيم في الحقيقة مصطنع أو غير حقيقي. إضافة لذلك، سوف نواجه في الفصول اللاحقة بعض الحالات التى تكون فيها الانحدارات مبنية على السببية، وبعضها ليس كذلك، والبعض الآخر لسنا متأكدين منه. الخلاصة العامة هنا هى أنك تحتاج إلى أن تكون حريصاً عند تفسير نتائج الانحدار باعتبارها تعكس السببية، ونفس الشيء ينطبق على نتائج الارتباط. الفهم العام والنظرية الاقتصادية سوف يساعدانك على تفسير كليهما.

3- إذا لم تفهم ذلك، قم بإنشاء المثال الرقمي الخاص بك، أى اختر أى قيم لكل من α وβ وX، ثم استخدم المعادلة $Y = \alpha + \beta X$ لحساب قيمة Y (وسمي هذه بـ Y الأصلية). والآن زد قيمة X بواحد صحيح، مع بقاء قيم α وβ دون تغيير ثم احسب قيمة Y الجديدة. لا تهم القيم التى اخترتها في الأساس لكل من α وβ وX لأنك ستجد أن قيمة Y الجديدة مطروحاً منها Y الأساسية (الأصلية) تساوي β تماماً. بمعنى آخر، أن β هى قياس الأثر على Y نتيجة لزيادة قيمة X بوحدة واحدة.

4- يحسب برنامج إكسل SSR,RSS,TSS في جدول تحليل التباين (ANOVA). العمود المعنون "SS" يشتمل على هذه الأرقام الثلاثة. في هذه المرحلة، قد لا تعلم ماذا تعنى

ANOVA، إلا أننا سوف نناقشها باختصار في الفصل السابع، تحت عنوان: "الانحدار مع وجود متغيرات صورية" Regression with Dummy variables.

5- يمكنك حساب اللوغاريثم الطبيعي لأي رقم في الإكسل باستخدام Formula bar على سبيل المثال، إذا كنت تريد أن تحسب اللوغريثم الطبيعي للرقم في الخلية D4 تحرك لـ formula bar واكتب "$= \ln(D4)$" واضغط على زر Enter.

6- هناك شيء أخر عن اللوغاريثمات الطبيعية يتمثل في أنها تنطبق فقط على الأرقام الموجبة. لذلك إذا كانت بياناتك تشتمل على أصفار أو أرقام سالبة، فإنه لايمكنك أخذ اللوغاريثمات الطبيعية لها (بمعنى آخر، سوف تعرض لك برامج الحاسب الآلي رسالة خطأ).

الفصل الخامس

الانحدار من منظور إحصائي

(Statistical Aspects of Regression)

الإحصاء هو مجال دراسة يعتمد على علم الرياضيات ونظرية الاحتمالات. ومع ذلك، وعلى الرغم من أن هذا الكتاب يفترض عدم معرفتك بهذه المواد، فإن الفهم الكامل للقضايا الإحصائية في نموذج الانحدار قد يتطلب وقتاً لدراستها بتركيز وعمق[1]. وبدلا من ذلك فسوف نقدم في هذا الكتاب ما يلي:

1- مناقشة ماهية المهمة التى تم من أجلها استخدام الأساليب الإحصائية في نموذج الانحدار.

2- توضيح كيفية تحليل الانحدار باستخدام هذه الأساليب الإحصائية وترجمة النتائج التى تم الحصول عليها.

3- تقديم بعض المفاهيم البيانية من أجل الحصول على فكرة مبسطة عن من أين تأتي النتائج الإحصائية ولماذا يتم ترجمة وتوضيح هذه النتائج بالطريقة التى تكون عليها (بمعنى آخر "كيف" و "لماذا" المتعلقة بالتحليلات الإحصائية).

وسوف نبدأ بالتركيز على التفريق الذي برز في الفصل السابق بين معاملات الانحدار α وβ ومقدرات المربعات الصغرى الاعتيادية لمعادلات الانحدار، $\hat{\alpha}$ و $\hat{\beta}$. تذكر أننا بدأنا الفصل الرابع بنموذج انحدار بالصيغة التالية:

$$Y_i = \alpha + \beta X_i + e_i$$

حيث i = 1،....، لعدد N من المشاهدات. وكما سبقت الإشارة إليه، فإن α وβ تقيس العلاقة بين Y و X. وقد أشرنا إلى أننا لانعلم ماهى هذه العلاقة، بمعنى آخر، ماهما α وβ على وجه التحديد. واشتققنا مايعرف بالمربعات الصغرى الاعتيادية، تقديرات (OLS) والتى عرفناها لاحقاً بـ $\hat{\alpha}$ و $\hat{\beta}$.

وركزنا على أن α وβ هما معاملان حقيقيان غير معلومان بينما $\hat{\alpha}$ و $\hat{\beta}$ ماهما إلا تقديرات (وبالتأكيد غالباً لا تساويان α وβ).

تقودنا هذه الاعتبارات للتساؤل عما إذا كان بإمكاننا قياس دقة هذه التقديرات. ولحسن الحظ، فإنه يمكننا استخدام الأساليب الإحصائية، خاصة وأن هذه الأساليب تمكننا من تقدير فترات ثقة لها، وتمكننا من إجراء اختبارات الفرضيات على معاملات الانحدار.

ولتقديم بعض الأمثلة، نقول إن المربعات الصغرى الاعتيادية تقدم نقطة تقدير لـ β (مثل $\hat{\beta} = 0.000842$ وهى قيمة مقدرة لـ β في انحدار الانحسار في مساحة الغابات على الكثافة السكانية في الفصل السابق). يمكنك التفكير في نقطة تقدير كأفضل تخمين عن ماهية قيمة β. فترات الثقة تقدم تقديرات فترية بما يسمح لنا صياغة عبارات تعكس حالة عدم التأكد التى قد نكون فيها حول القيمة الحقيقية لـ β (مثل "أننا نثق بأن قيمة β أكبر من 0.0006 وأقل من 0.0010"). يمكننا الحصول على فترات ثقة مختلفة تقابل مستويات مختلفة للثقة، على سبيل المثال، في حالة فترة ثقة 95%، يمكننا القول "أننا واثقون بنسبة 95% أن β تقع داخل نطاق فترة الثقة"؛ في حالة فترة ثقة بنسبة 90% يمكننا القول "أننا واثقون بنسبة 90% بأن β تقع في نطاق فترة الثقة" وهكذا. درجة الثقة التى لدينا في الفترة المختارة (على سبيل المثال 95%) يشار إليها كمستوى ثقة.

النشاط الرئيسي الآخر لباحث الاقتصاد القياسي هو اختبار الفرضيات. وكمثال للفرضية التى قد يريد الباحث اختبارها هى أن $\beta = 0$. فإذا كانت هذه

الفرضية صحيحة، فإن ذلك يعنى أن المتغير التفسيري ليس لديه قوة تفسيرية. الإجراءات المتعلقة باختبار الفرضية تسمح لنا بإجراء مثل هذه الاختبارات.

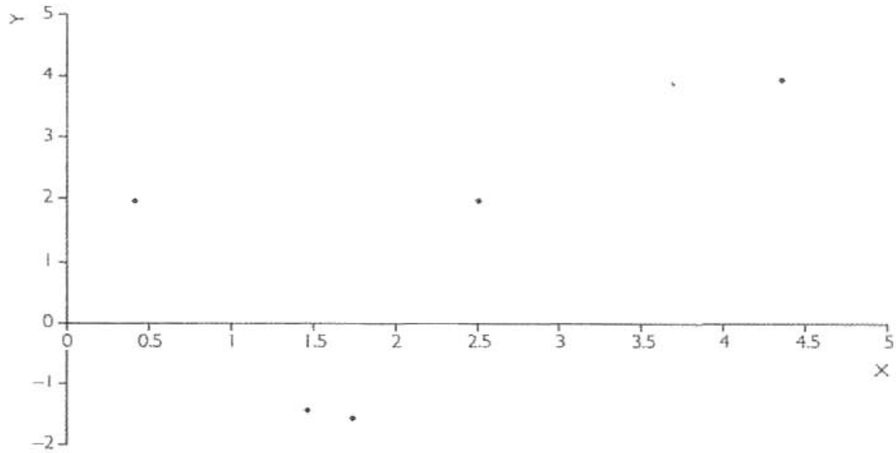
سوف يتم توضيح كل من فترات الثقة وإجراءات اختبار الفرضية بشكل أكثر فيما تبقي من هذا الفصل. وللأغراض التى قد تنشأ عرضاً، فإننا سوف نركز على β ، لأنها دائماً تكون أكثر أهمية من α في المسائل الاقتصادية. ومع ذلك، فإن كل الإجراءات التي سوف نناقشها بالنسبة لـ β تنطبق نفس الشيء على α.

ماهى العوامل التي تؤثر في دقة التقدير $\hat{\beta}$ ؟

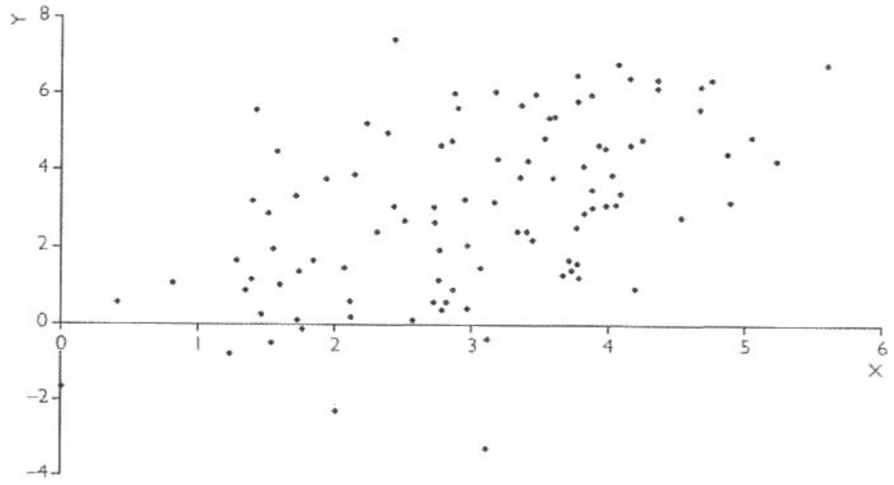
لقد قمنا بعمل محاكاة لأربع مجموعات من البيانات الصورية لكل من X و Y من نماذج الانحدار بافتراض أن $\alpha = 0$ و $\beta = 1$. وتوضح الأشكال (5-1 و 5-2 و 5-3 و 5-4) أشكال الانتشار لهذه المجموعات الأربع. كل هذه المجموعات من البيانات لديها نفس قيم المعاملات الحقيقية لكل من $\alpha = 0$ و $\beta = 1$ ، وعندما نقوم بتقدير النموذج من أى من هذه المجموعات الأربعة للبيانات نأمل أن نحصل على قيم كل من $\hat{\alpha}$ و $\hat{\beta}$ التي تكون مساوية تقريبا لـ 1 و صفر على التوالي، ومع ذلك، إذا تخيلت أنك تحاول رسم خط مستقيم (كما في حالة OLS) خلال نقاط شكل الانتشار، فلا يمكنك أن تتوقع أن تكون كل هذه الخطوط الأربعة متساوية بشكل تام.

ما مدى الثقة التى تشعر بها حول دقة الخط المستقيم الذي قمت برسمه؟ من البديهي أنك ترى أن الخط المرسوم للشكل رقم (5-3) هو الأكثر دقة. أى أن العلاقة التى يشكلها الخط المستقيم بين كل من X و Y "قد أصبحت أكثر ظهوراً" في الشكل رقم (5-3). وحتى إذا تم استخدام حدود مستقيمة وأنه يتوجب عليك

أن ترسم بيدك أفضل خط يربط بين النقاط في شكل الانتشار، فإنك سوف تجد أن القـاطع (α) كـان قريبـاً جـداً للـصفر وأن الميـل (β) كـان قريبـاً مـن الواحـد الصحيح. وبالمقابل، قد تكون أقل ثقة فيما يتعلق بدقة أفضل خط يربط بين النقاط والذي تم رسمه للأشكال (5-1 و 5-2 و 5-4).



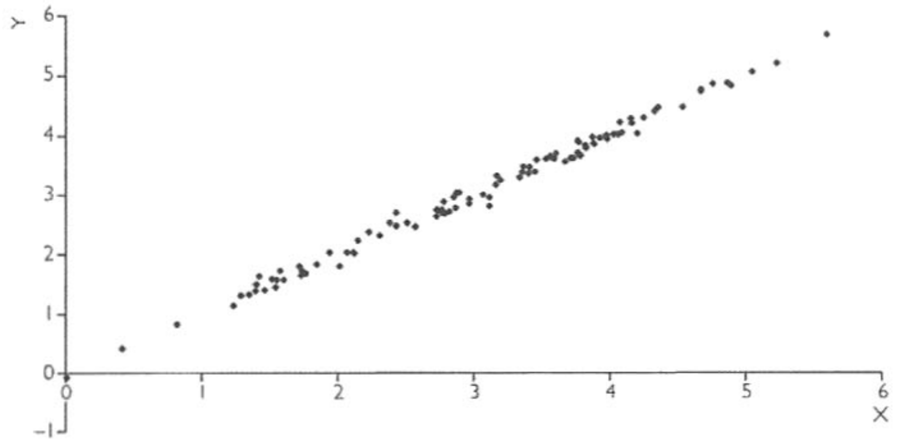
شكل رقم (5-1) حجم عينة صغير جدا



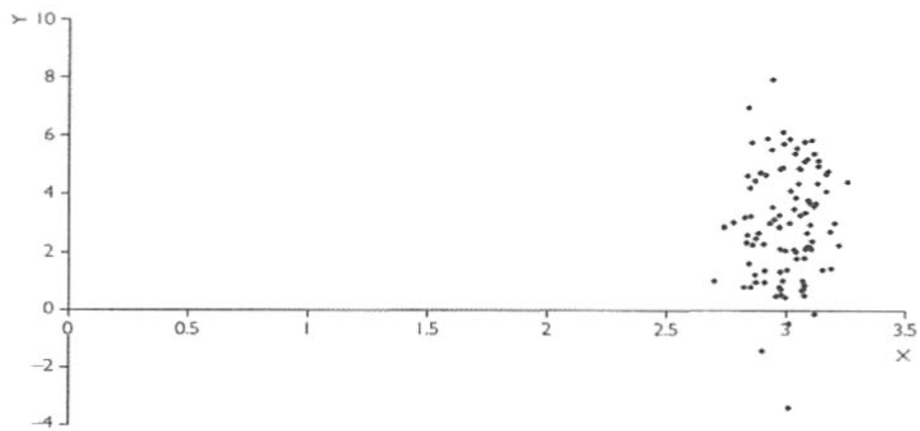
شكل رقم (5-2) حجم عينة كبير مع تباين ذي حد خطأ كبير

هذه الأشكال توضح ثلاثة عوامل رئيسة تؤثر في دقة تقديرات OLS وعلى عدم التأكد الذي يحيط بمعرفتنا حول القيمة الفعلية لـ β :

1- عندما يكون لدينا نقاط بيانات أكثر فإن ذلك يحسن دقة التقدير. ويمكن ملاحظة ذلك من خلال مقارنة الشكل رقم (5-1) (N=5) والشكل رقم (5-3) (N=100).



شكل رقم (5-3) حجم عينة كبير مع تباين ذي حد خطأ صغير



شكل رقم (5-4) مدى محدود لقيم X

2- إذا كان لديك أخطاء قليلة فإن ذلك يؤكد دقة التقدير. وبالمقابل، إذا كانت SSR صغيرة، أو إذا كان تباين الأخطاء صغيراً جدا، فإن دقة التقدير سوف تتحسن أكثر. ويمكن ملاحظة ذلك من خلال مقارنة الشكل رقم (5-2) (التباين مع حد خطأ كبير) مع الشكل رقم (5-3) (التباين مع حد خطأ صغير)[2].

3- إذا كان لديك انتشار كبير للقيم (بمعنى آخر تباين أكبر) للمتغير المستقل (X) فإن ذلك سوف يحسن من دقة التقدير. ويمكن ملاحظة ذلك من خلال مقارنة الشكل رقم (5-3) (تنتشر جميع قيم المتغير التفسيري على طول الخط من صفر وحتى الرقم 6) مع الشكل رقم (5-4) (تتجمع كل قيم المتغير التفسيري حول الرقم 3).

إنه من البديهي أن يكون لهذه العوامل الثلاثة تأثير معقول. وبالنسبة للعنصرين الأولين، فإنه من المعقول أن يكون لدينا إما بيانات كثيرة أو أخطاء صغيرة لزيادة دقة التقدير. العنصر الثالث قد يكون أقل تأثيراً، والمثال التالي سوف يساعدك على فهم ذلك.

افترض أنك ترغب في اختبار تأثير مستويات التعليم (X = سنوات الدراسة) على الدخل الذي يحصل عليه الأفراد (Y = الدخل). ولفهم طبيعة هذه العلاقة، فإنك سوف تحتاج إلى أن تذهب للشارع لمقابلة مختلف الأفراد (على سبيل المثال، الأفراد الذين ليس لديهم مؤهلات، الأفراد الذين حصلوا على الشهادة الثانوية، الأفراد الذين نالوا بعض التعليم فوق الثانوى كالتدريب المهنى، الأفراد الذين حصلوا على شهادات جامعية، الأفراد الحاصلون على الدكتوراه ... الخ). بمعنى آخر، إنك تريد مقابلة شريحة عريضة من السكان للحصول على أكبر قدر ممكن من هذه المستويات المختلفة للتعليم. في المفهوم الإحصائي، يعنى ذلك أنك تريد أن تحصل على تباين أكبر وإذا لم تتبع هذه الخطوات، وقمت مثلاً بمقابلة

حصرية للأفراد الحاصلين على الدكتوراه، فإنك سوف تحصل على صورة غير حقيقية لأثر التعليم على الدخل. وفى هذه الحالة فإنك لاتعرف ما إذا كانت العلاقة بين التعليم والدخل تكون إيجابية أم لا؟. على سبيل المثال، من غير جمع بيانات عن الأفراد الذين غادروا المدرسة في سن 16 سنة فإنك لن تعرف بالتأكيد بأنهم يحصلون على دخل أقل من حملة الدكتوراه.

الحصول على انتشار أكبر للقيم (بمعنى آخر، تباين أكبر) للمتغير المستقل X، هى ميزة مرغوبة في التحليل، بينما لايكون الأمر كذلك عند الحصول على انتشار أكبر للقيم (بمعنى آخر تباين أكبر) للخطأ العشوائي e.

حساب فترة الثقة للمعامل β :

عكست العوامل الثلاثة المشار إليها أعلاه في فترة تقدير شائعة الاستخدام للميل β : هى فترة الثقة. هذه الفترة (أو درجة القياس المعنوية) تعكس عدم التأكد الذي يحيط بدقة تقدير $\hat{\beta}$. فإذا كانت فترة الثقة صغيرة، فإنها تشير إلى دقة أكبر. وعلى العكس من ذلك، فإن فترة الثقة الكبيرة تشير إلى عدم تأكد عال حول القيمة الحقيقية لـ β's. وفي الكثير من الحالات يختار الباحثون أن يظهروا أو يقدموا فترة الثقة بالإضافة إلى (أو حتى بدلاً) نقطة التقدير الخاصة بالمربعات الصغرى الاعتيادية (OLS).

الصيغة الرياضية لفترة الثقة الخاصة بـ β هي[3]:

$$[\hat{\beta}-t_b s_b, \hat{\beta}+t_b s_b].$$

وهناك طريقة أخرى للتعبير عن المعادلة أعلاه، وهى من خلال القول بأنه ليس هناك مستوى مرتفع للثقة أى أن القيمة الحقيقية لـ β تتماشى مع المتطابقة (ingunliy) التالية:

$$\hat{\beta} - t_b S_b \leq \beta \leq \hat{\beta} + t_b S_b$$

المعادلة أعلاه استخدمت ثلاثة أرقام يجب حسابها، $\hat{\beta}$، t_b، S_b.

أولها: $\hat{\beta}$، والتي سبق أن ناقشناها بالتفصيل، أما t_b، S_b فقد لاتكون شاهدتهما من قبل. فترة الثقة يمكن حسابها تلقائياً في برامج الحاسب الآلي مثل برنامج إكسل. وبذلك يمكنك حساب فترات الثقة من غير معرفة أى من الصيغ الواردة أعلاه أو التعريف الدقيق لكل من t_b، S_b. ومن ناحية مبدئية فقط، يمكنك التفكير في $\hat{\beta}$، t_b، S_b كثلاثة أرقام يمكن حسابها بواسطة الحاسب الآلي.

ومع ذلك، فإنه من المفيد أن يتوافر لديك بعض المعرفة البديهية حول من أين تأتي فترة الثقة؟ لأن ذلك سوف يساعد في فهمنا لتلك النتائج.

فيما يلي سوف نناقش كلاً من الأرقام الثلاثة المطلوبة لحساب فترة الثقة، وربطها بتلك القضايا التى تتم إثارتها في الموضوع خلال نقاشنا حول العناصر التى تؤثر في دقة تقدير $\hat{\beta}$.

أولاً: غالباً ما يتم تضمين β في فترة الثقة (في الحقيقة، أنها سوف تكون في منتصفها تماماً).

ثانياً: S_b هى الانحراف المعياري لـ $\hat{\beta}$، وقد يكون مربكاً بعض الشيء أن S_b غالباً ما يشار إليها بأنها معامل الخطأ المعياري وهو ما يختلف عن الانحراف المعياري. في الفصل الثاني، أشرنا إلى الانحراف المعياري كمقياس للانتشار (بمعنى آخر، الانتشار، أو التغير) للمتغير. على سبيل المثال، يوضح الشكل رقم (2-2) المدرج التكراري للمتغير متوسط دخل الفرد من إجمالي الناتج المحلي باستخدام مجموعة البيانات المقطعية المتوافرة في ملف البيانات GDPPC.xls. في الفصل الثاني، أشرنا إلى أن الانحراف المعياري لمتوسط دخل الفرد من إجمالي الناتج المحلي كان مقياساً للاختلاف في مقدار دخل الفرد من إجمالي الناتج المحلي في الدول محل الدراسة. ومع ذلك قد يبدو من المناسب أن نتعامل مع $\hat{\beta}$ كمتغير بنفس طريقة أن متوسط دخل الفرد من الناتج المحلي هو متغير. بمعنى آخر، يمكننا حساب انحرافه المعياري واستخدامه كمقياس لعدم تأكدنا من دقة التقدير.

القيم الكبيرة لـ S_b تعني عدم تأكد عال. في هذه الحالة، قد تكون $\hat{\beta}$ تقدير غير دقيق جداً لـ β. وفى المقابل، فإن القيم الصغيرة لـ S_b تعنى عدم تأكد أقل. وإذا كان الأمر كذلك، فإن $\hat{\beta}$ سوف تُعد تقديراً مناسباً لـ β.

في الفصول الأخرى، قمنا بوضع المعادلات الرياضية في الملاحق، ورغم ذلك، فإنه يبقي مطلوباً أن يكون هناك قليل من الأساليب الرياضية حتى نتمكن من تحديد الارتباطات بين معادلة فترة الثقة والتوضيح البياني في الأشكال (5-1- 5-4).

ونقدم (إلا أننا لانشتق) المعادلة التالية للانحراف المعياري لـ $\hat{\beta}$:

$$S_b = \sqrt{\frac{SSR}{(N-2)\sum(X_i - \bar{X})^2}}$$

هذا التعبير الذي يقيس مستوى التغير أو عدم التأكد في قيمة $\hat{\beta}$، يعكس كل المسائل التي تم إبرازها من خلال نقاشنا للأشكال (5-1 و 5-2 و 5-3 و 5-4).

ومن النظر لمعادلة فترة الثقة يمكننا ملاحظة أنه كلما كانت S_b كبيرة، كانت فترة الثقة واسعة. وإذا ربطنا ذلك مع التحليل الدقيق لعناصر معادلة S_b، فإنه يمكننا القول أن:

1- S_b وعرض فترة الثقة، ترتبط عكسياً مع قيمة SSR (بمعنى آخر، أخطاء أكثر تبايناً (البواقي) تعني تقديراً أقل دقة).

2- S_b وعرض فترة الثقة، ترتبط عكسياً مع N (بمعنى آخر، مشاهدات أكثر تعنى تقديراً أكثر دقة).

3- S_b وتبعا لذلك، عرض فترة الثقة، ترتبط عكسياً مع $\sum(x_i - \bar{x})^2$ (بمعنى آخر، كثرة التغير في قيمة X تعني تقديراً أكثر دقة).

لاحظ أنه وكما ورد في الفصل الثاني، أن $\sum(x_i - \bar{x})^2$ هى عنصر رئيس للانحراف المعياري لـ X. وبصفة خاصة فإن القيم الكبيرة لهذا القياس تكون مرتبطة بقيم الانحراف المعياري الكبيرة لـ X.

ونؤكد هنا أن هذه العوامل الثلاثة (N، SSR والانحراف المعياري لـ X)، والتي تؤثر في عرض فترة الثقة، هي نفسها التي تمت مناقشتها أعلاه كمؤثرات على مدى دقة تقدير OLS للمعامل $\hat{\beta}$.

الرقم الثالث في المعادلة بالنسبة لفترة الثقة هو tb. ويبدو أنه من الصعوبة تقديم تخمين أكثر حول هذا الرقم بدون توافر بعض المعرفة أو الإلمام بالإحصاء.

وأما الذين لديهم إلمام بعلم الإحصاء، فيجب عليهم أن يضعوا في الاعتبار أن tb هي قيمة مأخوذه من الجداول الإحصائية لتوزيع t. ويستعرض الملحق رقم (1-5) تفاصيل إضافية حول tb. وعلى الرغم من ذلك، قد يساعدنا المثال التالي على فهم مدلولات هذا الرقم (t):

مثال: مراكز الاقتراع (الانتخاب):

قد تواجهك "نقطة تقديرات" أو أمر له علاقة بفترة الثقة (مستوى المعنوية) في مراكز الاقتراع السياسية، والتي يتم أخذها بانتظام قبل أسابيع أو شهور من إجراء الانتخابات. وهذه غالباً ما يقوم بإجرائها موظفون يتصلون هاتفياً ببضع مئات من الناخبين المحتملين ويسألونهم عن الحزب الذي سيقومون بانتخابه في يوم الانتخابات. وبافتراض أن الحزب A سيخوض الانتخابات، وأن التقارير الصحفية تشير إلى أن 43% من الذين جرى استطلاع آرائهم سوف يدعمون الحزب A. فإن هذه تعتبر نقطة التقدير الصحفية التي تشير إلى كيف سيصوت الناخبون يوم الاقتراع. وبالطبع، فإن الواقع يشير إلى أن النتيجة الحقيقية نادراً ما تكون مشابهة لما أشارت إليه استطلاعات الرأي إذا لم تخالفها تماماً. هذا الاختلاف يوضح نقطة ركزنا عليها سابقاً في هذا الفصل في سياق الحديث عن نموذج الانحدار ألا وهي نقطة التقدير (على سبيل المثال $\hat{\beta}$) التي نادراً ما تكون مشابهة للقيمة الحقيقية (على سبيل المثال β).

الصحف اليومية أيضاً تعرف أن مسوحاتها لن تكون دقيقة وغالباً ما يضيفون عبارات لتغطيتهم الصحفية مثل "هذه النتيجة صحيحة بزيادة أو نقصان بمعدل 2%". وعلى الرغم من أنهم لا يقولون ذلك صراحة فإنهم

يأخذون هذه النتيجة من فترة الثقة (التي غالباً ما تكون 95% مستوى معنوية)[4]. والعبارة المقابلة لذلك هي: "أننا واثقون بنسبة 95% من أن حزب A سوف يحصل على ما بين 41% و45% من أصوات الناخبين في يوم الاقتراع".

هذا المثال يقدم بعض البديهيات الإضافية حول ماهية فترات الثقة المعنوية. فإذا فهمت هذا المثال، يمكنك أيضاً معرفة أن مستويات الثقة المختلفة تعني أيضاً فترات ثقة مختلفة. وكمثال صغير على ذلك، انظر إلى مستوى الثقة المعنوية 100%. يمكننا أن نكون واثقين من أن الحزب A سوف يحصل على ما بين صفر إلى 100% من الأصوات في يوم الانتخابات. أي أن فترة ثقة 100% التي تمثل النسبة المئوية للأصوات التي حصل عليها الحزب A هي [0.100].

الآن انظر إلى النقيض من ذلك: ما هي درجة الثقة التي نكون عليها من أن الحزب A سوف يحصل على نسبة 43% شبه مؤكدة من الأصوات؟ وربما لا نكون واثقين جداً، وكما هو معروف، فإنه من النادر أن يكون هناك توافق بين استطلاعات الرأي ونتائج الانتخابات. ولهذا السبب، فإن فترة ثقة معنوية تكون نحو 43% (على سبيل المثال 42.9,43.1) يكون لديها مستوى ثقة منخفض جداً (ربما 10%).

لاحظ أنه كلما أردت أن تكون أكثر ثقة في فترة ثقتك المعنوية، فإنها تكون أكثر اتساعاً. على سبيل المثال، فترات الثقة بنسبة 99% تكون في الغالب أقل اتساعا من فترات الثقة بنسبة 95%. قيمة tb تتحكم في مستوى الثقة. فإذا كان مستوى الثقة عالياً (على سبيل المثال 99%) فإن tb يكون كبيراً، أما إذا كان مستوى الثقة منخفضاً، (على سبيل المثال 50%) فإن tb يكون صغيراً.

ولكي نعود للنظرية الإحصائية العامة للانحدار، يجب أن نركز (بدون زيادة التوضيح الوارد في المثال السابق) على ما يلي:

1- قيمة tb تتناقص مع قيمة N (بمعنى آخر، كلما كان لديك نقاط بيانات أكثر كانت فترة الثقة صغيرة).

2- تتزايد قيمة tb مع مستوى الثقة الذي تختاره.

غالباً ما يختار الباحثون مستوى ثقة عند 95% أو أيضا يمكن اختيار 99% أو 90% لفترات ثقة في بعض الأحيان). فهم بديهي مفيد (لكن غير دقيق على المستوى النظري) أن نقول حول مستوى الثقة 95% ما يلي:

(يمكن القول أن هناك احتمال 95% بأن القيمة الحقيقية لـ β تقع في حدود فترة الثقة 95%). التفسير الصحيح لهذه العبارة هو: (إذا استخدمت وبشكل متكرر، ولمجموعات بيانات مختلفة، معادلة فترة الثقة السابقة لحساب فترة ثقة 95% فإن القيمة الحقيقية لـ β ستكون داخل حدود فترة الثقة تلك). عبارات مماثلة يمكن استخدامها بالنسبة لمستويات الثقة 99%أو90%، وذلك بكل سهولة من خلال إحلال 99% أو 90% محل 95%. من الواضح أنه يمكن تفسير فترات الثقة بشكل مباشر (وسوف يتم توضيح ذلك أكثر من خلال الأمثلة اللاحقة في هذا الفصل).

الشرح السابق أعلاه، قصد منه تقديم بعض البديهيات والمحفزات للنظرية الإحصائية المتعلقة بفترات الثقة. وحتى إذا لم تفهم تلك المادة بشكل كامل، فإن فترة الثقة يمكن حسابها بسهولة بواسطة برامج الحاسب الآلي الجاهزة. على سبيل المثال، عندما تقوم بإجراء انحدار في برنامج إكسل فإنه يقوم بحساب فترة الثقة بشكل مباشر ويقدر الحدود الدنيا والعليا لدرجة الثقة المعنوية 95%. كذلك يُمكنك برنامج إكسل من تغيير مستوى الثقة، على سبيل المثال من 99% إلى 90%.

مثال: فترات الثقة لمجموعات البيانات في الأشكال (5-1 - 5-4):

تشتمل هذه الأشكال على أربع مجموعات بيانات مختلفة، كلها لديها $\alpha = 0$ و $\beta = 1$. تذكر بأن مجموعة البيانات المستخدمة في الشكل رقم (5-3) لديها بعض الخصائص المرغوبة جداً، بمعنى آخر لديها حجم عينة كبير، قيم انتشار كبيرة للمتغيرات المستقلة، وأخطاء صغيرة. هذه الخصائص تُعد مفقودة بدرجات متفاوتة في مجموعات البيانات الثلاثة الأخرى.

جدول رقم (5-1) تقديرات المربعات الصغرى الاعتيادية (OLS) وفترات الثقة

مجموعة البيانات	$\hat{\beta}$	فترة الثقة 90%	فترة الثقة 95%	فترة الثقة 99%
شكل رقم (5-1)	0.91	[-0.92.2075]	[-1.57,3.39]	[-3.64,5.47]
شكل رقم (5-2)	1.04	[0.75,1.32]	[0.70,1.38]	[0.59,1.49]
شكل رقم (5-3)	1.00	[0.99,1.01]	[0.99,1.02]	[0.98,1.03]
شكل رقم (5-4)	1.52	[-1.33,4.36]	[-1.88,4.91]	[-2.98,6.02]

يشتمل جدول رقم (5-1) على تقديرات المربعات الصغرى الاعتيادية (OLS) لـ β وفترات الثقة المعنوية لـ 90%،95% و99% لمجموعات البيانات الأربع المشار إليها أعلاه.

وبالنظر إلى الجدول يمكننا استخلاص النقاط التالية:

1- من خلال قراءة أي صف بيانات في الجدول أعلاه، نلاحظ أنه كلما كان مستوى الثقة عالياً، زادت فترة الثقة المعنوية. وأكبر مدى ثقة أو فترة ثقة هو فترة الثقة 99% لمجموعات البيانات في الشكل رقم (5-4). في هذه الحالة، إذا كنت تريد أن تكون واثقاً بنسبة 99%، يجب أن تقول أن β قد تقع بين القيمتين -2.98 و 6.02.

2- مجموعة البيانات في الشكل رقم (5-3)، وهي أكثر مجموعة تتميز بميزات مرغوبة مقارنة بكل مجموعات البيانات الأخرى – لها تقدير (OLS) يساوي 1.0 وهو يساوي القيمة الحقيقية لمنزلتين عشريتين ($\hat{\beta}=1.002577$).

3- مجموعة البيانات في الشكل رقم (5-3) لديها فترات ثقة أكثر ضيقاً مقارنة بالبيانات الأخرى في الأشكال (5-1و5-2 و5-4) وهذا له معنى وهو أننا نتوقع تقدير (OLS) باستخدام مجموعة البيانات في الشكل رقم (5-3) لتكون أكثر دقة من مجموعات البيانات الأخرى.

4- هناك اختلافات واضحة بنتائج مجموعات البيانات في الأشكال (5-1و5-2 و5-4). يحتوي الشكل رقم (5-2) على مجموعة بيانات من النوع الذي غالباً ما يكون مثالياً من الناحية التطبيقية (نادراً ما يجد الشخص مجموعة بيانات جيدة كما في الشكل رقم 5-3).

تتسم مجموعة البيانات هذه بأكثر الميزات المرغوبة، إلا أن معدل كثرة الأخطاء يميل إلى التوسط، مما يعكس الخطأ في القياس والنواقص التي تعتري النظرية الاقتصادية التي غالباً ما تحدث على أرض الواقع. وبالنسبة لمجموعة البيانات هذه فإن $\hat{\beta}=1.04$ وهي قيمة ليست بعيدة جداً من القيمة الحقيقية $\beta=1$. وفيما يخص هذه المجموعة من البيانات، يمكننا صياغة عبارة من النوع التالي "تقع قيمة β داخل فترة الثقة [0.70و1.38] بمستوى ثقة قدره 95%" أو "أننا واثقون بنسبة 99% أن β تقع بين 1.49 و 0.59".

تمرين رقم (5-1):

مجموعـات البيانـات المـستخدمة لحـساب الأشـكال (5-1و5-2 و5-3 و5-4) موجودة على الملفات FIG51.xls, FIG52.xls, FIG53.xls, FIG54.xls

(أ) احسب تقديرات (OLS) $\hat{\alpha}$ و$\hat{\beta}$ لهذه المجموعات الأربع من البيانات.

لأى مدى هي قريبة من الصفر والواحد الصحيح (تلك القيم التى استخدمناها لمحاكاة البيانات)؟

(ب) احسب فترات الثقة لـ $\hat{\alpha}$ لمجموعات البيانات الأربع. اختبر مدى علاقة عرض فترة الثقة بـ N وحالة التغير في قيمة الأخطاء.

(ج) احسب فترات الثقة بمستوى معنوية 99% و 90% لمجموعات البيانات. كيف تختلف عن فترات الثقة الخاصة بـ 95% في الفقرة (ب) أعلاه.

مثال: قياس انحدار انحسار مساحة الغابات على الكثافة السكانية

دعنا نعود لمجموعة بياناتنا حول الانحسار في مساحة الغابات (Y) والكثافة السكانية (X) في الملف (FOREST.xls). لقد رأينا في الفصل الأخير أن $\hat{\beta} = 0.000482$ ، بمعنى آخر، أن الأثر الحدي للكثافة السكانية على الانحسار في مساحة الغابات كان يعادل 0.000842. فترة الثقة 95% لهذا الأثر هى (0.00061,0.001075) ويؤكد ذلك وبدرجة ثقة عالية أن الأثر الحدي للسكان على الانحسار في مساحة الغابات أكبر من 0.00061 وأقل من 0.001075.

مثال: قياس انحدار مساحة الأرض على سعر المنزل

في الفصل السابق قمنا بقياس أثر مساحة الأرض = X على سعر بيع المنزل = Y باستخدام بيانات 546 منزلاً تم بيعها في مدينة ويندسور بكندا (انظر مجموعة البيانات في الملف HPRICE.xls) وبإجراء انحدار Y على X حصلنا على العلاقة المقدرة التالية:

$$Y=34,136+6.59X,$$

أو بمعنى آخر، أن $\hat{\alpha} = 34,136$ و $\hat{\beta} = 6.59$. يمكننا القول أن تقدير OLS للأثر الحدي لـ X على Y هو 6.59. وبذلك فإن توقعاتنا تقوم على أن زيادة مساحة الأرض بقدم مربع تزيد سعر المنزل بمقدار 6.59 دولار أمريكي.

فترة أو درجة الثقة بمستوى معنوية 95% بالنسبة لـ β هي (5.72,7.47). وعلى الرغم من أن أثر مساحة الأرض على سعر المنزل قد تم تقديره بـ 6.59 دولار أمريكي، إلا أننا غير متأكدين من أن هذا هو الرقم الصحيح. ومع ذلك، فإننا واثقون تماماً (بمعنى آخر أننا واثقون بدرجة 95%) أن فترة الثقة هذه سوف تكون كافية للمشتري المحتمل أو البائع لأخذ فكرة جيدة عن قيمة مساحة الأرض.

تمرين رقم (5-2):

يحتوى الملف ADVERT.xls بيانات عن Y (المبيعات السنوية) و X (مصاريف الإعلان) (بملايين الدولارات) لـ 84 شركة في الولايات المتحدة الأمريكية.

(أ) قم بإجراء انحدار لـ Y على X واحصل على فترات ثقة 95% بالنسبة لـ α و β.

(ب) اكتب عبارة توضح ماذا تعنى فترة ثقة 95% بالنسبة لـ β في شكل المدى المحتمل للقيم التى يؤثر بها المتغير التفسيري على المتغير التابع.

تمرين رقم (5-3):

يحتوى الملف ELECTRIC.xls على بيانات Y = تكاليف الإنتاج (بملايين الدولارات) و X = الإنتاج (بآلاف الكيلووات ساعة) لـ 123 شركة كهرباء في الولايات المتحدة الأمريكية.

كرر تمرين رقم (5-2) لهذه المجموعة من البيانات.

اختبار ما إذا كانت $\beta = 0$:

اختبار الفروض هو أسلوب آخر شائع الاستخدام بين الاقتصاديين القياسيين. وبالنسبة لفترات الثقة، فإننا لن نستخدم نظرية إحصائية تتضمن اختبار فرضيات. بدلاً عن ذلك، سوف نركز على التفاصيل العملية لكيفية إجراء اختبارات الفروض وشرح النتائج. اختبار الفرضيات التقليدي (الكلاسيكي) يتضمن تحديد للفرضيات التي يجب اختبارها. ويعرف ذلك بفرض العدم (null hypothesis) ويشار لذلك بالرمز H_0. وتتم مقارنتها بالفرض البديل (alternative hypothesis) ويشار إليه بالرمز H_1. اختبار الفرض الشائع هو هل $\beta = 0$؟. نظرياً، نقول إن هذا هو اختبار لـ $H_0: \beta = 0$: في مقابل $H_1: \beta \neq 0$.

لاحظ أنه إذا كانت $\beta = 0$ فإن X لا تظهر في نموذج الانحدار، أي أن المتغير التفسيري يفشل في تقديم أي قوة تفسيرية للمتغير التابع. إذا فكرت في أنواع الأسئلة التي يهتم بها الاقتصاديون (على سبيل المثال "هل يزيد التعليم دخل الفرد؟"، "هل تسهم إستراتيجية معينة في زيادة المبيعات؟"، "هل يسهم مشروع تدريب حكومي جديد في تخفيض البطالة؟"... الخ) فسوف ترى أنها جميعها على شكل "هل للمتغير التفسيري أثر في المتغير التابع؟" أو "هل $\beta = 0$ في انحدار Y على X؟". الغرض من اختبار فرض $\beta = 0$ هو الإجابة عن هذا السؤال.

النقطة الأولى التي يجب التأكيد عليها هى أن اختبارات الفروض وفترات الثقة يتقاربان بشكل كبير. وفي الحقيقة، فإن أحد الطرق التي يتم بها اختبار ما إذا كانت $\beta = 0$ هو النظر لفترة الثقة لـ β ومعرفة ما إذا كانت تحتوي على صفر. وإذا لم تكن كذلك، فإنه يمكننا ذكر بعض العبارات الإحصائية "أرفض الفرض القائل بأن $\beta = 0$ " أو أختم بـ "X لديها قوة تفسيرية كبيرة لـY" أو " β تختلف بشكل كبير عن الصفر" أو "تعتبر β معنوية إحصائياً". إذا كانت فترة الثقة تشتمل على الصفر فإننا نقوم بتغير كلمة رفض (reject) لكلمة قبول (accept) وعبارة "لديها قوة تفسيرية معنوية" إلى عبارة "ليس لديها قوة تفسيرية معنوية" وهكذا فإن طريقة فترة الثقة هذه لاختبار الفرض تساوي تماماً الطريقة العلمية لاختبار الفروض أو الفرضيات التي سوف تتم مناقشتها أدناه.

كما هو الحال بالنسبة لفترات الثقة التي تأتي بمستويات مختلفة للثقة (على سبيل المثال، 95% هي الخيار الغالب) تأتي اختبارات الفرضيات بمستويات معنوية مختلفة. فإذا استخدمت طريقة فترة الثقة لاختبار الفرض، فإن مستوى المعنوية هو 100% ناقصاً مستوى الثقة. أي أنه إذا كانت فترة ثقة 95% لا تشتمل على الصفر، فإنه يمكنك القول "أرفض الفرض القائل بأن $\beta = 0$ عند

5% مستوى معنوية" (بمعنى آخر، 100%-95%=5%). فإذا كنت قد استخدمت فترة ثقة 90% (ووجدت أنها لا تحتوي على الصفر) فإنه يمكنك القول: "أرفض الفرض القائل بأن $\beta = 0$ عند 10% مستوى معنوية".

الطريقة الأخرى لإجراء الفرض هي حساب اختبار إحصائي. في حالة اختبار ما إذا كانت $\beta = 0$، فإن الاختبار الإحصائي يُعرف بإحصاء-t (أو t-ratio أو-t stat). ويتم حسابه كما يلي:.

$$t = \frac{\hat{\beta}}{S_b}$$

القيم الكبيرة لـ t (الكبيرة بالمعنى المطلق) تشير إلى أن $B \neq 0$ بينما تشير القيم "الصغيرة" إلى أن $\beta = 0$. المفهوم البديهي للجملة السابقة هو: إذا كانت $\hat{\beta}$ كبيرة مقارنة بانحرافها المعياري S_b، فإنه يمكننا أن نخلص إلى أن β تختلف معنوياً عن الصفر. ويبرز السؤال هنا، ماذا نعني بـ "الكبيرة" و "الصغيرة"؟. في المعنى الإحصائي النظرى، يعد الاختبار الإحصائي كبيراً أو صغيراً مقارنة مع "القيمة الحرجة" التي تؤخذ من الجداول الإحصائية الخاصة بـ "Student-t distribution". ويستعرض الملحق رقم (5-1) كيفية القيام بذلك. لحسن الحظ، أننا لانحتاج لبذل جهد كبير للنظر في الجداول الإحصائية لأن معظم برامج الحاسب الآلي الجاهزة مثل برنامج إكسل تقدم لنا ما يعرف بقيمة الاحتمال (p-value) بطريقة تلقائية. قيمة الاحتمال توفر قياساً مباشراً لما إذا كانت t "كبيرة" أو "صغيرة". ويعد تفسير قيمة الاحتمال على أنه مقياس الاحتمال أن $\beta = 0$ تفسيراً مفيداً من الناحية العملية (ولكنه ليس دقيقاً علمياً). فإذا كانت قيمة الاحتمال صغيرة، فإن افتراض أن $\beta = 0$ ليس صحيحاً. ووفقاً لذلك فإنه يمكن القول أن:

1- إذا كانت قيمة الاحتمال أقل من 5% مستوى معنوية (عادة يتم كتابتها 0.05 بواسطة الحاسب الآلي) فإن قيمة t "كبيرة" ونتوصل إلى أن $\beta \neq 0$.

2- إذا كانت قيمة الاحتمال أكبر من 5%، فإن قيمة t تكون "صغيرة" وبذلك نتوصل إلى أن $\beta = 0$.

الاختبار السابق استخدم 5% مستوى معنوية. ومع ذلك، إذا كنا نريد استبدال الرقم 5% في المثال أعلاه بـ 1% (بمعنى آخر رفض $\beta = 0$ إذا كانت قيمة الاحتمال أقل من 1%) فإننا سوف نجري اختبار الفرض عند مستوى 1% مستوى معنوية.

وفي المقابل، تجدر الإشارة إلى أننا نركز على اختبار أن $\beta = 0$ لأنه مهم من ناحية، ولأنه عادة مايتم حسابه بواسطة برامج الحاسب الآلي من ناحية أخرى. ويمكنك استخدامه من غير أن تكون ملماً إلماماً كافياً بالنظرية الإحصائية التى يقوم عليها. وبالرغم من ذلك، فإنك إذا أردت اختبار فروض أخرى (على سبيل المثال، $H_0 : \beta = 1$ أو فروض تتضمن معاملات عديدة في حالة الانحدار المتعدد في الفصل القادم) فإنك تحتاج إلى معرفة إحصائية أكثر مما تم تغطيته هنا (انظر الملحق رقم (5-1) لتفاصيل أكثر). الهيكل العام لاختبار الفرض غالباً ما يكون على الشكل الذى تم توضيحه أعلاه، أى أنه:

(1) يحدد الفرض الذي سيتم اختباره.

(2) يحسب اختباراً إحصائياً.

(3) يقارن الاختبار الإحصائي بالقيمة الحرجة.

الخطوة الأولى من هذه الخطوات الثلاث تعد سهلة، إلا أن الخطوتين الثانية والثالثة أكثر صعوبة. وعلى وجه الخصوص، فإن الحصول على اختبار

إحصائي لاختبارات فرض أكثر تعقيداً تتطلب حسابات إضافية تتعدى مسألة إجراء انحدار.

ويتطلب أيضاً الحصول على القيمة الحرجة استخدام الجداول الإحصائية. لذلك إذا أردت إجراء اختبارات فروض أكثر تعقيداً فإنه يجب عليك الرجوع إلى مرجع أساسي عن أساسيات الإحصاء أو الاقتصاد القياسي (انظر الملاحظة رقم 1 لهذا الفصل لبعض المقترحات).

وكملخص عملي، خذ في الاعتبار أن أساليب الانحدار تقدم المعلومات التالية حول β :

1- $\hat{\beta}$ ، هى قياس المربعات الصغرى الاعتيادية (OLS) للنقطة، أو أفضل تخمين لقيمة β .

2- فترة الثقة 95%، التى تعطي فترة نكون فيها واثقين بنسبة 95% أين تقع قيمة β .

3- الانحراف المعياري (أو الخطأ المعياري) لـ $\hat{\beta}$ ، S_b ، هو قياس لمدى صحة $\hat{\beta}$. S_b هو أيضاً عنصر رئيس في المعادلة الرياضية لفترة الثقة والاختبار الإحصائي لاختبار الفرض العدم $\beta = 0$.

4- إحصائية الاختبار t ، لاختبار أن $\beta = 0$.

5- قيمة الاحتمال (p-value) لاختبار أن $\beta = 0$.

هذه العناصر الخمسة ($\hat{\beta}$ ، فترة الثقة، S_b ، t وقيمة الاحتمال) عادة ما تظهر في صف واحد في برامج الحاسب الآلي الجاهزة مثل برنامج إكسل. في

الواقع العملي، فإن أكثرها أهمية هى $\hat{\beta}$ ، فترة الثقة، وقيمة الاحتمال. وعادة يمكنك شرح نتائجك القياسية من غير الرجوع المباشر لكل من t و S_b. وتوضح الأمثلة التالية كيف يتم حساب نتائج الانحدار وكيفية تفسيرها.

مثال: انحدار الانحسار في مساحة الغابات على الكثافة السكانية. (تتمة ما ورد في الصفحات السابقة)

إذا قمت بإجراء انحدار لـ Y (الانحسار في مساحة الغابات) على X (الكثافة السكانية)، باستخدام برنامج إكسل، فإنه يمكن الحصول على النتائج التى تظهر في جدول رقم (5-2) أدناه (برامج الحاسب الآلي الأخرى سوف تقدم نتائج مشابهة لها):

جدول رقم (5-2) انحدار انحسار مساحة الغابات على الكثافة السكانية

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة %95 (Lower 95%)	الحد الأعلى لفترة الثقة %95 (Upper 95%)
القاطع (Intercept)	0.599965	0.112318	5.341646	1.15E-06	0.375837	0.824093
المتغير X	0.000842	0.000117	7.227937	5.5E-10	0.00081	0.001075

الصف المعنون بالقاطع (Intercept) يحتوي على نتائج لـ α ، والصف المعنون "المتغير X" "X-variable" يشتمل على نتائج لـ β . وسوف نركز نقاشنا على الصف الأخير. العمود المعنون بالمعامل "coefficient" يبرز تقدير المربعات الصغرى الاعتيادية لـ β ، و كما رأينا سابقاً، $\hat{\beta} = 0.000842$ ، مما يشير إلى أن زيادة الكثافة السكانية بمقدار شخص واحد لكل هكتار يرتبط بزيادة في معدلات الانحسار في مساحة الغابات بنسبة 0.000842%. الأعمدة المعنونة بـ "الحد الأدنى لفترة الثقة 95% و الحد الأعلى لفترة الثقة 95%" تعطي الحدود الدنيا والعليا لفترة الثقة 95% لهذه المجموعة من البيانات، وكما سبق نقاشه، فإن فترة الثقة 95% هي (0.00061 و 0.001075). لذلك، فإننا واثقون بنسبة 95% من أن الأثر الحدي للكثافة السكانية على الانحسار في مساحة الغابات هو بين 0.001075 و 0.00061%.

الأعمدة المعنونة الخطأ المعياري وإحصائية t تشير إلى أن $s_b = 0.000117$ و t=7.227937. هذه الأرقام ليست ضرورية لإجراء اختبار الفرض أن $\beta = 0$ عندما تكون قيمة الاحتمال (p-value) معطاة. ولأغراض كثيرة يمكننا تجاهل هذين العمودين[5] .

يمكن إجراء اختبار الفرض أن $\beta = 0$ بطريقتين متساويتين. أولهما، يمكننا إيجاد فترة الثقة 95% بالنسبة لـ β (0.001075 و 0.00067). ولأن فترة الثقة هذه لا تشتمل على الصفر، فإنه يمكننا رفض الفرض القائل بأن $\beta = 0$ عند مستوى معنوية 5%. بمعنى آخر، هنالك إثبات قوي للفرض بأن $\beta \neq 0$ وأن الكثافة السكانية لها قوة كبيرة في تفسير الانحسار في مساحة الغابات.

ثانيهما، يمكننا النظر إلى قيمة الاحتمال والتي هي $5.5X10^{-10}$ [6]، وأقل كثيراً من 0.05. وهذا معناه أنه يمكننا رفض الفرض القائل بأن الكثافة السكانية ليس لديها أي أثر في انحسار مساحة الغابات عند مستوى معنوية 5%. بمعنى آخر، لدينا إثبات قوي بأن الكثافة السكانية تؤثر بالفعل في معدلات الانحسار في مساحة الغابات.

تمرين رقم (5-4):

مستخدماً جدول رقم (5-2) (قم بإجراء انحدار مستخدماً مجموعة البيانات الموجودة في الملف FOREST.XLS)، قم باختبار الفرض القائل بأن $\alpha = 0$.

تمرين رقم (5-5):

إضافة لـ Y= معدل الانحسار في مساحة الغابات، فإن ملف البيانات FOREST.XLS يحتوي أيضاً على بيانات عن W= نسبة الزيادة في الأراضي الزراعية (معنونة "Crop ch") و Z= نسبة التغير في أراضي المراعي (المعنونة "pasture ch").

(أ) قم بإجراء انحدار لـ Y على W واشرح النتائج التي توصلت إليها. هل يمكنك رفض الفرض القائل بأن التوسع في الأراضي الزراعية ليس له أثر في معدلات الانحسار في مساحة الغابات؟.

(ب) قم بإجراء انحدار لـ Y على Z واشرح النتائج التي توصلت إليها. هل يمكنك رفض الفرض القائل بأن أراضي المراعي ليس لها أثر في معدلات الانحسار في مساحة الغابات؟

تمرين رقم (5-6):

باستخدام مجموعة البيانات في الملفات FIG51.xls, FIG52.xls, FIG53.xls و FIG54.xls

(أ) اختبر ما إذا كانت $\beta = 0$ مستخدماً طريقة فترة الثقة لكل مجموعة من مجموعات البيانات الأربع.

(ب) اختبر ما إذا كانت $\beta = 0$ مستخدماً طريقة قيمة الاحتمال ومجموعة البيانات الأربع. استخدم مستوى المعنوية 5%.

(جـ) أعد ما قمت به في كل من (أ) و (ب) أعلاه بالنسبة لـ α.

(د) أعد ما قمت به في كل من (أ)، (ب) و (ج) مستخدماً مستوى معنوية 1%.

(هـ) هل إجاباتك ذات معنى مفيد في ضوء النقاش الذي أجريناه في هذا الفصل عن العوامل التي تؤثر في دقة تقديرات المربعات الصغرى الاعتيادية (OLS)؟

مثال:

وجدنا سابقاً أن فترة الثقة 95% في انحدار Y (سعر المنزل) على X (مساحة الأرض) وهي (5.27,7.47). ولأن فترة الثقة لا تحتوي على الصفر، فإنه يمكننا رفض الفرض القائل بأن $\beta = 0$ عند مستوى معنوية 5%. بالفعل فإن مساحة الأرض لها أثر إحصائي مهم في أسعار المنازل.

وفي المقابل، قيمة الاحتمال تساوي $6.77X10^{-42}$، وهي تقل كثيراً عن 0.05. وكما في السابق، يمكننا رفض الفرض القائل بأن $\beta = 0$ عند مستوى معنوية 5%. لاحظ أيضاً أنه ما دام $6.77X10^{-42}$ يقل عن 0.01 فإنه يمكننا رفض الفرض القائل بأن $\beta = 0$ عند مستوى معنوية 1%. هذا إثبات قوي بالفعل على أن مساحة الأرض تؤثر فعلاً في أسعار المنازل.

تمرين رقم (5-7):

لقد استخدمنا ملف ADVERT.XLS في السابق. تذكر بأنه يحتوي على بيانات حول المبيعات وتكاليف الإعلان لعدد84 شركة. قم بإجراء انحدار مستخدماً هذه البيانات وناقش نتائجك شفاهة كذلك ضعها في شكل تقرير. ضمن ذلك نقاشاً حول الأثر الحدي لتكاليف الإعلان على المبيعات والنقاش حول ما إذا كان هذا الأثر الحدي ذا معنوية إحصائية.

اختبار الفرض الذي يتضمن R^2 : (إحصائية F)

معظم برامج الحاسب الآلي الجاهزة التي تحسب الانحدار مثل برنامج إكسل، تقدم أيضاً نتائج عن اختبار الفرض $H_o: R^2 = 0$. التعريف بـ R^2 والشرح المتعلق بها قد ورد في الفصل السابق. وبالأخذ في الاعتبار أن R^2 هي مقياس للكيفية التي يتوافق بها خط الانحدار مع البيانات بشكل جيد، أو بصورة متساوية هي مقياس لنسبية التغير في Y التي يمكن شرحها من خلال X. إذا كانت $R^2 = 0$ فإن X ليس لديها أى قوة تفسيرية لـ Y. اختبار الفرض بأن $R^2 = 0$ يمكن توضيحه كاختبار حول ما إذا كان الانحدار يوضح أي شيء على وجه العموم. في حالة الانحدار البسيط فإن هذا الاختبار يساوي اختبار $\beta = 0$.

في الفصل التالي سوف نناقش حالة الانحدار المتعدد (حيث يكون هناك العديد من المتغيرات التفسيرية)، ويكون اختبار الفرض مختلفاً. وكمقدمة للنقاش الذي سوف نجريه في الفصل التالي، ضع في الاعتبار أن القيمة $R^2 = 0$ سوف يتم استخدامها لاختبار ما إذا كانت كل المتغيرات التفسيرية لديها أى قوة تفسيرية

للمتغير التابع. وفي المقابل فإنه سيتم استخدام إحصائية t-statistic لـ $\beta = 0$ لمعرفة ما إذا كان متغيراً مستقلاً واحداً لديه قوة تفسيرية.

ولا تختلف عملية اختبار أن $R^2 = 0$ عن المنطق المشار إليه أعلاه. أي أن برامج الحاسب الآلي تحسب اختباراً إحصائياً ويجب عليك مقارنته مع القيمة الحرجة. وفي المقابل، فإنه يمكن حساب قيمة الاحتمال التي تعطي مباشرةً قياساً لاختبار الفرض العدم بأن $R^2 = 0$، في مقابل الفرض البديل، $R^2 \neq 0$. معظم البرامج الإحصائية سوف تحسب تلقائياً قيمة الاحتمال، وإذا كان الأمر كذلك، فإنك لا تحتاج لمعرفة الشكل الدقيق للاختبار الإحصائي أو كيف تستخدم الجداول الإحصائية للحصول على القيمة الحرجة. ولإكمال الصورة تبعاً لذلك، فإننا نقدم إحصائية ف، F-statistic،[7] والتي يتم حسابها على النحو التالي:-

$$F = \frac{(N-2)R^2}{1-R^2}$$

هذه المعادلة يتم حسابها تلقائياً بواسطة برنامج إكسل وتتم تسميتها بقيمة "F".

كما في السابق، فإن القيم "الكبيرة" للاختبار الإحصائي تشير إلى أن $R \neq 0$ بينما القيم "الصغيرة" تشير إلى $R^2 = 0$. بالنسبة لاختبار $\beta = 0$، فإننا نستخدم قيمة الاحتمال لنقرر ماهو "كبير" وما هو "صغير" (بمعنى آخر ما إذا كانت R^2 تختلف بشكل كبير عن الصفر أم لا). لاحظ مع ذلك أن برنامج إكسل يرجع لقيمة الاحتمال لهذا الاختبار بحسب "معنوية F" – "Significance F". يتم إجراء الاختبار وفقا للخطوات التالية:

1- إذا كانت معنوية F أقل من 5% (بمعنى آخر 0.05) فإننا نقول $R^2 \neq 0$.

2- إذا كانت معنوية F أكبر من 5% (بمعنى آخر 0.05) فإننا نقول $R^2 = 0$.

الطريقة السابقة تقدم اختباراً إحصائياً بمستوى معنوية 5%. لإجراء اختبار عند مستوى معنوية 1%، فما عليك إلا استبدال 5% (0.05) بـ 1% (0.01) في العبارات السابقة. يمكن حساب مستويات معنوية أخرى (على سبيل المثال 10%) بنفس الطريقة.

تستخدم بعض برامج الحاسب توضيحاً يختلف قليلاً عن ما يفعله برنامج إكسل. على سبيل المثال، Micro Fit تعنون F-statistic بـ "F-stat" وتضع قيمة الاحتمال بين قوسين بعد F.

مثال: انحدار انحسار مساحة الغابات على الكثافة السكانية (تتمة ما ورد في الصفحات السابقة):

في حالة مجموعة بيانات الانحسار في مساحة الغابات على الكثافة السكانية، فإن (F=52.24308). هل هذه القيمة "كبيرة"؟ إذا قلت نعم فإنك على حق، ما دامت معنوية F = 5.5×10^{-10}، وهي تقل عن 0.05. في ضوء هذه النتيجة يمكننا القول أن الكثافة السكانية لديها قوة تفسيرية لـ Y. عادة يمكننا القول بأن " R^2 تختلف معنوياً عن الصفر عند مستوى معنوية 5%"، X كان لديها قوة تفسيرية معنوية إحصائياً لـ Y" أو أن "الانحدار معنوى". لاحظ أن معنوية F تساوي قيمة الاحتمال في الاختبار الخاص بـ $\beta = 0$، مبرزاً مساواة هذين الاختبارين في حالة الانحدار البسيط.

تمرين رقم (5-8):

استخدم مجموعات البيانات في الملفات FIG52.xls, FIG53.xls, FIG51.xls و FIG54.xls

اختبر ما إذا كانت $R^2 = 0$ لكل مجموعة من المجموعات الأربع. قارن نتائجك بنتائج تمرين رقم (5-6).

مثال: تكلفة إنتاج الكهرباء

لقد استخدمنا الملف ELECTRIC.xls في الفصل الرابع، نظراً لاحتوائه على بيانات عن Y= تكاليف الإنتاج و X= الإنتاج في 123 شركة كهرباء. فإذا قمنا بإجراء انحدار لـ Y على X باستخدام برنامج إكسل فإننا سوف نحصل على النتائج الموضحة في الجدول رقم (5-3) أدناه.

جدول رقم (5-3): انحدار تكلفة الإنتاج على حجم الإنتاج

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
القاطع(Intercept)	2.186583	1.879484	1.163395	0.246958	-1.534354	5.90752
المتغير X	0.004789	0.000132	36.37623	5.36E-67	0.004528	0.005049

إضافة لذلك فإن $R^2 = 0.916218$. قيمة الاحتمال لقياس $R^2 = 0$ (التي يعنونها برنامج إكسل بـ "Significance F" "F المعنوية" تساوي 5.36E-67.

في هذا السياق، وتلخيصاً للمادة التي تمت مناقشتها في الفصلين الرابع والخامس، يجدر أن نوضح هنا كيف يمكن كتابة النتائج المشار إليها أعلاه في تقرير رسمي. التقرير سوف يتضمن عرضاً للمادة الإحصائية في الجدول أعلاه يلي ذلك ملخص يناقش التفسير والمنطق الاقتصادي وراء التحليل والنتائج الإحصائية التي تم الحصول عليها في ضوء تلك البديهية. وقد يكون شكل التقرير على النحو التالي:

يعرض الجدول رقم (3-5) نتائج عن انحدار المربعات الصغرى الاعتيادية (OLS) باستخدام مجموعة بيانات صناعة الكهرباء في الولايات المتحدة الأمريكية [8]. وما دمنا نهتم بالكيفية التى تؤثر بها خيارات الإنتاج المختلفة للمؤسسات على تكاليف إنتاجها، سوف نختار تكاليف الإنتاج كمتغير تابع والإنتاج كمتغير تفسيري.

هذا الجدول يوضح أن المعامل المقدر للإنتاج هو 0.004789، ويشير إلى أن المؤسسات الكهربائية ذات مستويات الإنتاج العالية لها تكاليف إنتاج عالية. فزيادة الإنتاج بألف كيلووات للساعة سوف تزيد التكاليف بمقدار 4789 دولاراً أمريكياً. يمكن ملاحظة أن الأثر الحدي للإنتاج على التكاليف هو أثر ذو معنوية إحصائية قوية، ما دامت القيمة الاحتمالية صغيرة جداً (صغيرة بشكل كبير، أى، أقل من 1%).

اختبار فترة الثقة 95% يوضح أنه يمكننا أن نكون واثقين كلياً أن زيادة الإنتاج بألف كيلو وات للساعة يترتب عليه زيادة في التكاليف بـ 4528 دولاراً أمريكياً على أقل تقدير و 5049 دولاراً أمريكياً على أكثر تقدير. اختبار R^2 يعيد إلى الأذهان أن الإنتاج يقدم الجزء الأكبر من التفسير حول لماذا تختلف التكاليف بين المنشآت الكهربائية. وعلى هذا فإنه يمكن تفسير 92% من حجم التغير في تكاليف الإنتاج على مستوى المؤسسات من خلال مستويات الإنتاج المختلفة. القيمة الاحتمالية لإحصائية F -F (ststistic) تقل كثيراً عن 1%، مما يشير إلى معنوية R^2 عند مستوى 1%.

ملخص الفصل:

1- تعتمد دقة تقديرات المربعات الصغرى الاعتيادية على عدد المشاهدات، وعلى مدى المتغير التفسيري ومدى الأخطاء.

2- تقدم فترة الثقة تقدير فترة لـ β (بمعنى آخر، الفترة التى تكون فيها على ثقة من أن β تقع فيها). يمكن حسابها في معظم برامج الحاسب الآلي الجاهزة.

3- عرض فترة الثقة يعتمد على نفس العوامل التى تؤثر في دقة تقديرات المربعات الصغرى الاعتيادية. إضافة لذلك، يعتمد عرض فترة الثقة على مستوى الثقة (بمعنى آخر، درجة الثقة التى تريد الحصول عليها في تقديرات فترة الثقة).

4- يمكن استخدام اختبار الفرض حول ما إذا كانت $\beta = 0$ لمعرفة ما إذا كان المتغير التفسيري يجب أن يضمن للانحدار. القيمة الاحتمالية التى يتم حسابها تلقائياً في برامج الحاسب الآلي الإحصائية، تُعد مقياساً لكيفية انطباق هذا الفرض.

5- إذا كانت القيمة الاحتمالية لاختبار الفرض حول ما إذا كانت $\beta = 0$ تقل عن 0.05 فإنه يمكنك رفض الفرض عند مستوى المعنوية 5%. وبذلك يمكنك القول بأن X تضمن في الانحدار.

6- إذا كانت القيمة الاحتمالية لاختبار الفرض ما إذا كانت $\beta = 0$ أكبر من 0.05 فإنه لايمكنك رفض الفرض عند مستوى معنوية 5%. وبذلك لايمكنك القول بأن X موجودة داخل الانحدار.

7- اختبار الفرض ما إذا كانت $R^2 = 0$ يمكن استخدامه لاختبار ما إذا كان الانحدار يساعد في تفسير المتغير التابع. القيمة الاحتمالية لهذا الاختبار يمكن حسابها تلقائياً في معظم تغيرات برامج الحاسب الآلي الإحصائية ويمكن استخدامها بنفس الطريقة التى تم توضيحها في النقاط 5 و 6 أعلاه.

ملحق رقم (5-1): استخدام الجداول الإحصائية لاختبار ما إذا كانت $\beta = 0$

القيمة الاحتمالية هى كل ما تريد أن تعرفه من أجل اختبار الفرض القائل بأن $\beta = 0$. معظم برامج الحاسب الآلي (مثل Excel, Microfit or Shazam) تقدم القيمة الاحتمالية تلقائياً. ومع ذلك، إذا لم يكن لديك مثل هذه البرامج أو أنها فقط صفحة تُظهر إحصاء t، وليس القيمة الاحتمالية، فإنه يكون من المفيد أن تعرف كيف يمكن اختبار الفرض باستخدام الجداول الإحصائية. في الواقع أن أى كتاب مرجعي في الإحصاء أو الاقتصاد القياسي سوف يعرض الطريقة بشكل أكثر تفصيلاً وسوف يقدم أيضاً الجدول الإحصائي الضروري لتقوم بذلك. هنا سوف نقدم نقاشاً مختصراً وواضحاً للغاية، وهو ما ينطبق على الحالة التي يكون فيها حجم العينة، N ، كبيراً.

لاحظ أن اختبار الفرض يتضمن المقارنة بين القيمة المحسوبة في الاختبار الإحصائي مع القيمة المجدولة (القيمة الحرجة). فإذا كانت قيمة الاختبار الإحصائي كبيرة (في قيمتها المطلقة) مقارنة بالقيمة الحرجة، فإن الفرض يُرفض. هنا الاختبار الإحصائي هو اختبار t (t-stat) الذي تم تقديمه في هذا الفصل. هذا يجب مقارنته بالقيمة الحرجة المأخوذة من الجدول الإحصائي لـ t (Student-t). وهو يشير إلى أن القيمة الحرجة هى تماماً ما سميناه t_b في نقاشنا عن فترات الثقة. وإذا كانت N كبيرة وأنت تستخدم مستوى معنوية 5%، فإن $t_b = 1.96$. وتبعاً لذلك فإنه يمكننا الاسترشاد بالقاعدة العامة التالية:

إذا كانت إحصائية t أكبر من 1.96 بالقيمة المطلقة (بمعنى آخر $|t| > 1.96$)، أرفض الفرض $\beta = 0$ عند مستوى معنوية 5%. إذا كانت إحصائية t أقل من 1.96 بالقيمة المطلقة، أقبل الفرض $\beta = 0$ عند مستوى معنوية 5%.

إذا تم رفض الفرض $\beta = 0$ ، يمكننا القول أن "X" ذات معنوية إحصائياً أو أن "X" تقدم قوة تفسيرية ذات معنوية إحصائية لـ "Y".

وتعتبر هذه القاعدة عامة إذا كان حجم العينة كبيراً.

من المعروف أن القيمة الحرجة تساوى 1.96 إذا كان حجم العينة إلى ما لا نهاية.

ومع ذلك، وحتى في حالات أحجام العينات المتوسطة فإنك سوف تحصل على قيم حرجة مشابهة. على سبيل المثال، إذا كانت N=120 فإن قيمتها الحرجة هى 1.98. وإذا كانت N=40 فإن قيمتها الحرجة هى 2.02. حتى حجم العينة الصغير جداً مثل N=20 فإنها تحصل على قيمة حرجة 2.09 والتى لاتختلف كثيرا عن 1.96.

ومع ذلك، يجب عليك أن تكون حذراً عند استخدام هذا الاختبار اليدوي إذا كانت N صغيرة جداً أو إذا كانت إحصائية t قريبة جداً لـ 2.00. إذا رجعت للأمثلة المضمنة في هذا الفصل، فإنك ترى أن الطريقة المشار إليها هنا تعمل جيداً. على سبيل المثال، في مثالنا المعنون "تكلفة الإنتاج في منشأة صناعة الكهرباء"، نجد أن قيمة إحصاء t لاختبار الفرض ما إذا كانت $\beta = 0$ كانت 36.4، التى تعتبر كبيرة جداً مقارنة 1.96.

وبذلك يمكننا القول أن الإنتاج متغير تفسيري ذو معنوية إحصائية لتكاليف الإنتاج. في هذا المثال (وكل الأمثلة الأخرى)، فإن كلاً من طريقتي القيمة الاحتمالية وفترة الثقة تقودان لنفس النتيجة التى وصفتها الطريقة في هذا الملحق.

النقاش السابق يتعلق بمستوى معنوية 5%. القيمة الحرجة للعينة الكبيرة بالنسبة لمستوى معنوية 10% هو 1.65. وبالنسبة لمستوى معنوية 10% فإن هذه القيمة هى 2.58.

ويعتبر أكثر اختبارات الفروض شيوعاً على الإطلاق هو $H_0: \beta = 0$. وباستخدام الأساليب التى ورد شرحها في هذا الملحق، فإنه يمكننا تعميم هذا الفرض بشكل طفيف إلى ما يلي: $H_0: \beta = c$ ، حيث c هى رقم قد لا يكون مساوياً للصفر (على سبيل المثال c=1). في هذه الحالة، تتغير إحصائية الاختبار (test statistic) بشكل طفيف، إلا أن القيمة الحرجة هى نفسها كما هو الحال لاختبار $\beta = 0$. وهنا تصبح إحصائية الاختبار كما يلي:

$$t = \frac{\hat{\beta} - c}{Sb}$$

هذه المعادلة لايتم حسابها أو عرضها بصورة آلية من خلال برامج الحاسب الآلي الإحصائية، إلا أنه يمكن حسابها بسهولة في برنامج إكسل أو الآلة الحاسبة. أى أنه يتم حساب كل من $\hat{\beta}$ و S_b بواسطة برامج الحاسب الآلي، وعليك إيجاد c، معتمداً على الفرض الذي ترغب في اختباره. ويمكن جمع هذه الأرقام الثلاثة باستخدام المعادلة أعلاه لتعطيك قيمة إحصائية الاختبار. إذا كانت هذه القيمة أكبر من 1.96 بالقيمة المطلقة، فإنه يمكنك القول أن $\beta \neq c$ عند مستوى معنوية 5%. هذا، ويجب مراعاة عدم استخدام هذه القاعدة العامة أو القاعدة التقريبية إذا كان حجم العينة صغيراً جداً.

ملاحظات ختامية:

1- كما ورد ذكره، فإن إحدى المقدمات الجيدة في أساسيات الإحصاء هو كتاب:

Introductory Statistics for Business and Economics

تأليف Thomas Wonacott and Ronald Wonnacott

(Fourth edition, John Wiley and Sons, 1990).

وفي أساسيات الاقتصاد القياسي كتاب Undergraduate Econometrics

تأليف R.Carter Hill, William Griffith and George Judge

(Second edition, John Willy&Sons 2000)

2- إذا كان لديك مشكلة في فهم هذه النقطة، ارسم خطاً مستقيماً بقاطع = 0.5 وميل يساوي 1 في الشكلين (5-2 و 5-3) وانظر لبعض البواقي الناتجة (المنشأة بحسب ما ورد في الشكل رقم 4-1). سوف ترى أن معظم البواقي في الشكل رقم (5-2) أكبر بكثير (في القيمة المطلقة) مقارنة بنظيراتها في الشكل رقم (5- 3). سوف ينتج ذلك عن SSR كبيرة (انظر المعادلة في الفصل الرابع)، ولأن البواقى والأخطاء هى أشياء متشابهة جداً، فإن هناك تبايناً كبيراً للأخطاء (انظر المعادلة الخاصة بالانحراف المعيارى للمتغير في الإحصاءات الوصفية في القسم الثاني من الفصل الثانى، وتذكر أن التباين ما هو إلا مربع الانحراف المعياري).

3- المعلومة القائلة بأن "المتغير W يقع بين a and b" و "W أكبر من أو تساوى a وأقل من أو تساوى b" يُعبر عنها رياضياً على هذا النحو "تقع W في الفترة [a,b]" وسوف نستخدم هذا التعبير الرياضي من وقت لأخر في هذا الكتاب.

4- اختيار فترة الثقة 95% هو الأكثر شيوعاً، وحيثما لايتم تحديد مستوى الثقة فإنه يمكنك افتراض أنه 95%.

5- في أمثلة هذا الكتاب لانستخدم أبداً S_b ونادراً ما نستخدم t. وكمرجعية مستقبلية فإن المواضع التي نستخدم فيها t هما اختبارا ديكي فولير (Dickey-Fuller) وإنجل- جرانجر (Engle-Granger) اللذان سوف يتم مناقشتهما في الفصلين 9 و 10.

6- تذكر أن 5.5E-10 هو الطريقة التي تكتب بها معظم برامج الحاسب الآلي 5.5×10^{-10} والتي يمكن كتابتها أيضا كالتالي 0.000.000.000.55.

7- نظرياً، تعد إحصائية F هى الوحيدة في كل مجموعة إحصائيات الاختبار التى تأخذ قيماً حرجة من ما يسمى توزيع F (F-distribution). يقدم الملحق رقم (1-11) نقاشاً إضافياً لذلك.

8- في كتابة أي تقرير، يتم في العادة تخصيص قسم منفصل ملحق بالتقرير لمناقشة البيانات المستخدمة وتفاصيلها. انظر الملحق (أ) حول إعداد مشروع بحث (قياسي) تطبيقي لنقاش أوسع حول كيف ينظم محتويات مشروع بحث اعتيادي.

الفصل السادس

الانحدار المتعـدد

(Multiple Regression)

تضمن النقاش في الفصل الخامس حول الانحدار البسيط متغيرين: المتغير التابع، Y، والمتغير التفسيري، X. كما ناقشنا في بداية الفصل الرابع، فإن معظم التحليلات في علم الاقتصاد تتضمن العديد من المتغيرات. يقوم الانحدار المتعدد بتوسيع الانحدار البسيط ليشمل العديد من المتغيرات التفسيرية. ومادام ذلك هو الأداة الشائعة الاستخدام في علم الاقتصاد القياسي (التطبيقي)، فإن هذا الفصل يجب أن ينظر إليه باعتباره من أهم أجزاء هذا الكتاب. لحسن الحظ أن معظم الفهم البديهي والأساليب الإحصائية للانحدار المتعدد تتشابه إلى حد كبير مع تلك المتعلقة بالانحدار البسيط.

ويمكن إجمال العناصر الأساسية للفصلين الرابع والخامس كما يلي:

1- استخدام الرسم البياني للفهم البديهي لأساليب الانحدار، مثل ملاءمة (توافق) الخط المستقيم خلال شكل الانتشار.

2- تقديم معامل الانحدار كمقياس للأثر الحدي.

3- وصف لتقدير المربعات الصغرى الاعتيادية (OLS) كأفضل خط ملاءمة (بالنسبة لتصغير مجموع البواقي المربعة) من خلال شكل الانتشار.

4- تقديم R^2 كمقياس لملاءمة نموذج الانحدار.

5- تقديم أساليب إحصائية مثل فترات الثقة واختبارات الفروض.

مع بعض الاستثناءات (تم إيضاحها أدناه)، هذه العناصر الخمسة تتفق مع نموذج الانحدار المتعدد. إذا وجدت صعوبة في تذكر البديهيات مدار البحث أو النواحي الإحصائية للانحدار فيجب عليك الرجوع للفصلين الرابع والخامس. هذا الفصل يغطي العناصر الخمسة في حالة الانحدار المتعدد باختصار شديد، مقدماً ملخصاً للتشابهات والاختلافات مقارنة بنموذج الانحدار البسيط. في معظم الفصل سيتضمن النقاش الأمثلة التي ستوضح كيف يمكن تفسير نتائج الانحدار المتعدد.

مثال: تفسير أسعار المنازل:

تركز الكثير من الأبحاث في مجال الاقتصاد الجزئي التطبيقي ومجال التسويق على كيفية تسعير السلع. أحد الطرق الشائعة في هذا الشأن تتضمن بناء نموذج يكون فيه سعر السلعة معتمداً على مواصفات تلك السلعة، وتحتوي مجموعة بيانات الملف HPRICE.Xls على بيانات تطبيقية عن ذلك، والذي يسمى أسلوب التسعير المرتبط بالإشباع (hedonic) لسوق العقار. وقد استخدمنا جزءاً من مجموعة البيانات هذه في الفصول السابقة، وبالأخذ في الاعتبار أنها تشتمل على بيانات عن 546 منزلاً تم بيعها في مدينة ويندسور بكندا. متغيرنا التابع، Y، هو سعر البيع للمنزل بالدولار الكندي، ومساحة الأرض هى متغيرنا التفسيري.

بالطبع فإن سعر المنزل يتأثر بأكثر من مساحة الأرض، وأى محاولة جادة لتوضيح محددات أسعار المنازل يجب أن تشتمل على أكثر من متغير تفسيري بخلاف مساحة الأرض. فى هذا الفصل، سوف نركز على أربعة متغيرات تفسيرية وهى:

$X_1 =$ مساحة الأرض (بالأقدام المربعة).

$X_2 =$ عدد غرف النوم.

$X_3 =$ عدد الحمامات.

$X_4 =$ عدد الأدوار (بخلاف القبو).

تشتمل مجموعة البيانات HPRICE.xls أيضاً على متغيرات تفسيرية أخرى سوف يتم استخدامها في الفصول التالية وفى التمارين اللاحقة.

تمرين رقم (6-1):

أ- ارسم أشكال انتشار مستخدماً المتغيرات التفسيرية الأربعة في مثال تسعير المنازل كل واحد على حدة (بمعنى آخر، ارسم X_1, Y ، ثم ارسم X_2, Y ، ... الخ).

ب- قم بإجراء انحدار بسيط مستخدماً المتغيرات التفسيرية كل واحد على حدة (بمعنى آخر، انحدار Y على X_1 ، ثم انحدار Y على X_2 ... الخ).

ت- دوِّن ملاحظاتك على العلاقات التي توصلت إليها في كل من أ و ب أعلاه.

الانحدار أفضل خط ملاءمة:

كما لاحظنا في الفصل الرابع، فإنه يمكن النظر إلى نموذج الانحدار البسيط بوصفه أسلوباً يهدف إلى رسم خط خلال شكل الانتشار. وحيث إن الانحدار المتعدد يعني وجود أكثر من متغيرين (على سبيل المثال، X_1 ، X_2 ، X_3 ، و X_4)، فإنه لايمكننا رسم شكل انتشار في رسم بياني ذي بُعدين، والذي يكون فيه أحد المتغيرين مرسوماً على المحور العمودي والآخر على المحور الأفقي. فإن نفس الخط الذى يتلاءم يكون موجوداً (ويمكن توضيحه فقط إذا كان بإمكاننا، بطريقة أو بأخرى، رسم رسوم بيانية ذات أبعاد متعددة). على سبيل المثال، إذا كان لدينا ثلاثة متغيرات تفسيرية، فإنه يمكننا إيضاح كيف يتضمن الانحدار المتعدد رسم خط ملاءمة خلال رسم بياني رباعي الأبعاد، تكون فيه Y مرسومة على محور واحد، X_1 على المحور الثاني، X_2 على المحور الثالث، و X_3 على المحور الرابع. الرسم البياني الرباعي لن يكون منتظماً إلى حد كبير وعملياً لايمكن رسمه (بمعنى آخر، كيف يبدو رسماً بيانياً من أربعة أبعاد؟).

تقدير المربعات الصغرى الاعتيادية لنموذج الانحدار المتعدد:

يمكن كتابة نموذج الانحدار المتعدد بوجود K متغيرات تفسيرية على النحو التالي [1]:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e.$$

بدلا عن تقدير α و β فقط، فإنه الآن لدينا $\alpha, \beta_1, \beta_2, ..., \beta_k$.

ومع ذلك، فإن طريقة الحصول على التقديرات لكل هذه المعاملات هى تماماً نفس تلك الخاصة بنموذج الانحدار البسيط. وعلى ذلك، فإننا نُعرف مجموع البواقى المربعة (SSR) على النحو التالي:

$$SSR = \sum (Y_i - \hat{\alpha} - \hat{\beta}_1 X_{1_i} - ...... - \hat{\beta}_k X_{k_i})^2$$

حيث X_{1_i} هى المشاهدة i_{th} على المتغير التفسيري الأول (بالنسبة لـ i=1,...,N مشاهدات، على سبيل المثال مساحة المنزل i حيث i=1,..........,546).

وقد جرى تعريف المتغيرات التفسيرية الأخرى بصورة مشابهة. ويمكن الحصول على تقديرات المربعات الصغرى الاعتيادية (OLS) (باعتبارها تقدم أفضل خط ملاءمة) من خلال اختيار قيم $(\hat{\alpha}, \hat{\beta}_1, ..., \hat{\beta}_k)$ التى تخفض SSR إلى حدها الأدنى. نظرياً، يُعد ذلك مسألة رياضية بحتة [2]، والمعادلات الناتجة معقدة لا نتطرق لنقاشها هنا [3]. مع العلم بأن برامج الحاسب الآلي مثل إكسل سوف تقوم بحساب تقديرات المربعات الصغرى الاعتيادية $(\hat{\alpha}, \hat{\beta}_1, ..., \hat{\beta}_k)$ بصورة تلقائية.

الانحدار المتعدد من منظور إحصائي:

كما هو معلوم، فإن الجوانب الإحصائية للانحدار المتعدد تتشابه مع حالة الانحدار البسيط (انظر الفصل الخامس). وأيضاً، ما زالت R^2 هي مقياس التوافق وتحسب بنفس الطريقة. لاحظ مع ذلك أنها تستخدم كمقياس للقوة التفسيرية لكل المتغيرات التفسيرية في نفس الوقت بدلاً عن أخذها كمتغير تفسيري وحيد في نموذج الانحدار البسيط. وبنفس الطريقة، فإن إحصائية F (F- statistic) لاختبار أن $R^2 = 0$ له معادلة تختلف قليلاً، حيث إن (N-2 تم استبدالها بـ N-K-1) إلا أنها بالأهمية نفسها وأنت تنظر إلى معنوية F (Significance F) في نتائج برنامج إكسل. فإذا وجدنا أن $R^2 \neq 0$، فحينها يمكننا القول أن: المتغيرات التفسيرية فى الانحدار، تساعد على تفسير التغير في المتغير التابع، بينما إذا وجدنا أن $R^2 = 0$، يمكننا القول أن المتغيرات التفسيرية ليست ذات معنوية ولا تقدم أى قوة تفسيرية للمتغير التابع.

المعادلات الرياضية العامة التي تستخدم لحساب فترات الثقة لمعاملات الانحدار ولاختبار ما إذا كانت المعالم تساوي صفراً هي نفسها المستخدمة في الفصل السابق. على الرغم من أن الأرقام الفعلية التى تشتمل عليها هذه المعادلات (على سبيل المثال، S_b) قد تم حسابها بطريقة أكثر تعقيداً. ومع ذلك، تبقي البديهية العملية كما هى. بمعنى آخر، أن فترة ثقة 95% سوف تقدم تقدير فترة مثل ما يمكنك القول "أننى واثق بنسبة 95% أن المعامل يقع فى فترة الثقة 95%". في برنامج إكسل، العمودان "الأدنى 95%" و "الأعلى 95%" مازالا الحدين الأدنى والأعلى لفترة الثقة 95%. إذا كان الرقم فى عمود "قيمة الاحتمال" أقل من 0.05 يمكننا القول أن المتغير التفسيري المعني أو تحت الدراسة ذو معنوية إحصائية عند مستوى 5%. يجدر بالذكر أن هناك الآن قيمة

احتمالية وفترة ثقة مرتبطة مع كل من المعاملات $(\beta_1,....,\beta_k)$ أكثر من أنها مرتبطة مع β واحدة فى نموذج الانحدار البسيط. ومع ذلك، فمن وجهة نظر الباحث الذى يأمل فى تفسير مخرجات الحاسب الآلى لاستخدامها في تقرير ما، فإن النواحى الإحصائية للانحدار المتعدد مشابهة لتلك الخاصة بالانحدار البسيط[4].

تفسير تقديرات المربعات الصغرى الاعتيادية:

توجد في تفسير تقديرات المربعات الصغرى الاعتيادية بعض الاختلافات المهمة بين حالتى الانحدار البسيط والانحدار المتعدد. هذا القسم سوف يقدم بعض الطرق لكيفية تفسير المعاملات في نموذج الانحدار المتعدد. وقبل أن نبدأ، لابد من تحديد واضح للمصطلحات التي سوف نستخدمها.

عندما نتحدث عن الخاصية التى ترتبط بالمعاملات فإننا نعرف المعامل بـ β_j (بمعنى آخر، المعامل المتغير التفسيري j_{th}، حيث j يمكن أن تكون أى رقم بين 1وK). عندما نود أن نتحدث عن معامل معين سوف نعطي الرقم المحدد لـ j (على سبيل المثال في حالة β_1 ، $j=1$، وهو معامل المتغير التفسيري الأول).

في حالة الانحدار البسيط رأينا كيف يمكن تفسير β كأثر حدي (بمعنى آخر، كمقياس للأثر الذى يحدثه التغير في X على Y أو كمقياس لتأثير X على Y). في الانحدار المتعدد، ما زال من الممكن تفسير β_j كأثر حدي، إلا أن ذلك يتم بطريقة تختلف قليلاً.

وعلى هذا فإن β_j هي الأثر الحدي لـ X_j على Y، مع بقاء كل المتغيرات الأخرى ثابتة (وتعني باللغة اللاتينية القديمة ceteris paribus، وهى عبارة

شائعة الاستخدام في علم الاقتصاد). هذه العبارة ذات أهمية كبيرة للتفسير الصحيح لنتائج الانحدار. ولهذا السبب، سوف نأخذ بعض الوقت لنشرح بدقة ماذا نعني بها، وذلك من خلال مثالنا عن أسعار المنازل.

مثال: تفسير أسعار المنازل (تتمة ما ورد في الصفحات السابقة):

يشتمل الجدول رقم (6-1) على نتائج من انحدار Y= سعر البيع على X_1 = مساحة الأرض، X_2 = عدد غرف النوم، X_3 = عدد الحمامات و X_4 = عدد الأدوار. صُمم الجدول لتحاكي نتائجه نتائج برنامج إكسل، وكذلك تعرض برامج الحاسب الآلي الأخرى نتائجها بطريقة مشابهة.

العمود الأول يوضح المتغيرات التفسيرية. في هذا المثال هناك أربعة متغيرات تفسيرية (إضافة للقاطع). كل صف يحتوى على نفس المعلومات كما في جدول نموذج الانحدار البسيط (بمعنى آخر، تقدير المربعات الصغرى الاعتيادية للمعامل المناسب يليه انحرافه المعياري، إحصائية t، قيمة الاحتمال لاختبار ما إذا كانت $\beta_j = 0$ والحدود الدنيا والعليا لفترة الثقة 95% للمعامل).

وكما تم أعلاه، فإن هذه النتائج الإحصائية الآن متاحة لكل معامل وجميعها سوف تكون مختلفة (على سبيل المثال، قيمة الاحتمال لاختبار ما إذا كانت $\beta_1 = 0$ سوف تكون مختلفة عن قيمة الاحتمال لاختبار ما إذا كانت $\beta_3 = 0$).

وباستخدام المعلومات المتوافرة في جدول رقم (6-1) يمكننا كتابة معادلة الانحدار الملائمة (المتوافقة) كما يلي:

$$\hat{Y} = -4009.55 + 5.43\,X_1 + 2824.61\,X_2 + 17105.17\,X_3 + 7634.90\,X_4$$

وكمثال، خذ المعامل للمتغير التفسيري الأول، مساحة الأرض. فإنه يمكن ملاحظة أن $\hat{\beta}_1 = 5.43$

يوجد أدناه طرق (مشابهة جداً) تنص على ماذا تعنى هذه القيمة.

1- "زيادة قدم مربع واحد لمساحة الأرض سوف يضيف 5.43 دولار على سعر المنزل، مع بقاء العوامل الأخرى كما هى".

جدول رقم (6-1) انحدار أسعار المنازل على مساحة المنزل، عدد غرف النوم، عدد الحمامات وعدد الأدوار*.

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
القاطع (Intercept)	-4009.5500	3603.109	-1.1128	0.266287	-11087.3	3068.248
X_1	5.4291737	0.369250	14.70325	2.05E-41	4.703835	6.154513
X_2	2824.61379	1214.808	2.325153	0.020433	438.2961	5210.931
X_3	17105.1745	1734.434	9.862107	3.29E-21	13698.12	20512.22
X_4	7634.897	1007.974	7.574494	1.57E-13	5654.874	9614.92

* لاحظ أنه فى هذا الجدول كما هو الحال في الجداول الأخرى، قد كتبنا الأرقام كما يعرضها برنامج إكسل. أى أننا استخدمنا المنازل العشرية لأقصي حد ممكن واستخدمنا الحرف "E" كرمز لقوة الرفع (exponents).

وعندما تكتب تقريراً فمن المحتمل أن ترغب في استخدام منازل عشرية قليلة، واستبدال 1.57E-13 بـ 1.57×10^{-13} إضافة لذلك، فإن $R^2 = 0.54$ وقيمة الاحتمال لاختبار أن $R^2 = 0$ (التي تعرف بمعنوية F (Significance F) في برنامج إكسل) هي 1.18E-88.

2- "إذا أخذنا منازل بنفس عدد غرف النوم، والحمامات وعدد الأدوار، فإن زيادة مساحة الأرض بمقدار قدم مربع سوف تزيد سعر المنزل بمقدار 5.43 دولار".

3- إذا قمنا بإجراء مقارنة بين المنازل التي لها نفس عدد غرف النوم والحمامات وعدد الأدوار، فإن المنازل ذات المساحة الكبيرة سوف تكون أعلى سعراً. وعلى هذا، فإن أي قدم مربع إضافي في مساحة المنزل سوف يترتب عليها زيادة بمقدار 5.43 دولار في السعر".

من المفيد التوسع في الدافع الذي يقف وراء التعبيرين السابقين. فلا يمكننا القول بسهولة أن "المنازل ذات المساحات الكبيرة ذات قيمة أعلى" ولأن الوضع ليس كذلك دائماً (على سبيل المثال، بعض المنازل الجميلة المشيدة على مساحات صغيرة سوف تكون قيمتها أعلى من المنازل غير الجميلة ومشيدة على مساحات كبيرة). ومع ذلك، فإنه يمكننا القول أنه "إذا أخذنا المنازل التي تختلف في المساحة، إلا أنها تتقارب في النواحي الأخرى، فإن المنازل ذات المساحات الكبيرة تكون ذات قيمة أكبر". التعبيران المذكوران أعلاه، يتضمنان بشكل صريح المزايا "إلا أنها تتقارب في النواحي الأخرى". ويلاحظ أننا لم نضمن هذه المواصفات في الفصل الرابع. وبالمقارنة، دعنا نأخذ $\hat{\beta}_2$ (معامل عدد غرف النوم)، و2842.61، ويمكن التعبير عنه كما يلي:

1- "المنازل التي فيها غرفة نوم إضافية تفوق نظيرتها التي لا تحوي غرفة نوم إضافية بما قيمته 2842.61 دولاراً، مع بقاء العوامل الأخرى كما هي".

2- "إذا أخذنا المنازل المتقاربة (ذات المزايا المتقاربة) (على سبيل المثال، ذات المساحات 5000 قدم مربع، حمامين ودورين)، فإن المنازل المكونة من ثلاث غرف نوم تفوق في قيمتها المنازل ذات غرفتي النوم بما قيمته 2842.61 دولاراً".

هنالك العديد من الطرق للتعبير عن تفسيرات هذه المعاملات. إلا أن العبارة العامة التى نريد صياغتها هي كالتالي: في حالة الانحدار البسيط يمكننا القول أن "β تقيس أثر X على Y"، في الانحدار المتعدد نقول إن "β_j تقيس أثر X_j على Y مع بقاء جميع المتغيرات التفسيرية الأخرى كما هى". العبارات المذكورة أعلاه ليست إلا طرقاً مختلفة للقول أن "كل المتغيرات التفسيرية الأخرى تبقي كما هى".

ويمكن تفسير المعاملات للمتغيرات التفسيرية الأخرى بطرق مشابهة. على سبيل المثال، $\hat{\beta}_3 = 17105.174$. وبالكلمات، يمكننا القول أن "المنازل ذات الحمام الإضافي يزيد سعرها بما قيمته 17.105.17 دولاراً مع بقاء العوامل الأخرى ثابتة أو كما هى".

ولأن $\hat{\beta}_4 = 7634.897$، فإنه يمكننا القول "إذا قارنا المنازل المتشابهة فى كل النواحى الأخرى، فإن ذات الدور الإضافي سوف يزيد سعرها بما قيمته 7634.90 دولاراً".

خذ في اعتبارك أنه عند مناقشة الجوانب الإحصائية لمعاملات الانحدار، فإن فترة الثقة وقيمة الاحتمال هما أكثر الأرقام أهمية، ويمكن تفسيرهما بنفس طريقة الانحدار البسيط. على سبيل المثال، ما دامت قيمة الاحتمال لكل المتغيرات التفسيرية (ماعدا القاطع) أقل من 0.05 يمكننا القول بأن "المعاملات $\beta_1, \beta_2, \beta_3, \beta_4$ ذات معنوية إحصائية عند مستوى 5%"، أو بنفس القدر أنه "يمكننا رفض الفروض الأربعة المنفصلة القائلة بأن أياً من هذه المعاملات هو صفر عند مستوى معنوية 5%".

وكمثال آخر، دعنا نأخذ فترة الثقة 95% بالنسبة لـ β_2 والتي هي (438.2761-5210.931) هذه المعلومات يمكن التعبير عنها كما يلي: "على الرغم من أن نقطة تقديرنا تشير إلى أن الأثر الحدي لعدد غرف النوم على أسعار المنازل هو 2842.61 دولاراً، فإن هذا التقدير ليس دقيقاً. فترة الثقة 95% تشير إلى أنه يمكننا أن نكون واثقين فقط بأن هذا الأثر الحدي (الإضافي) يقع بين 438.28 دولاراً و 5210.93 دولاراً". وبالمقارنة، فإن فترة الثقة بالنسبة لـ β_4 هي (5654.874-9614.92) ويمكننا القول: "أننا واثقون بنسبة 95% من أن الأثر الحدي لعدد الأدوار على سعر المنزل يقع بين 5654.87 دولاراً و 9614.92 دولاراً".

اختبار الفرض حول ما إذا كانت $R^2 = 0$ تحمل قيمة احتمالية تقل كثيراً عن 5%، مما يشير إلى أن X_1, X_2, X_3, X_4 لها قيمة تفسيرية ذات معنوية إحصائية للمتغير التابع. في الحقيقة أن الاختلافات في مساحة الأرض وعدد غرف النوم والحمامات وعدد الأدوار تصل إلى 54% من نسبة التغير في أسعار المنازل.

أوجه القصور في استخدام الانحدار البسيط في سياق الانحدار المتعدد:

لإبراز الفرق بين الانحدار البسيط والانحدار المتعدد، سوف نقوم بإجراء انحدار بسيط لـ Y= سعر المبيعات على X_2 = عدد غرف النوم. يشتمل جدول رقم (6-2) على نتائج من هذا الانحدار. بما أن $\hat{\beta}=13,269.98$ في هذا الانحدار البسيط، فإنه يمكننا صياغة عبارات من نوع "الأثر الحدي لعدد غرف النوم على أسعار المنازل هو 13269.98 دولاراً" أو "المنازل التي بها غرفة نوم إضافية تكون تكلفتها أكثر بما قيمته 13269.98 دولاراً".

يجب عليك مقارنة هذه العبارة بالعبارة السابقة لها. بالنسبة للانحدار البسيط فإننا حذفنا شرط بقاء العوامل الأخرى ثابتة أو كما هي المضمنة في جزء من عبارة: إذا أخذنا المنازل المتقاربة (على سبيل المثال، تلك المكونة من مساحات 5000 قدم مربع، حمامين ودورين...).

لاحظ أيضاً أن معامل عدد غرف النوم في الانحدار البسيط يكون أعلى من نظيره في الانحدار المتعدد. لماذا يكون هذا الوضع؟ للإجابة عن هذا السؤال، أولاً تخيل أن صديقة في مدينة ويندسور ترغب في بناء غرفة نوم إضافية في منزلها وسألتك، بصفتك اقتصادياً، كم تضيف غرفة نوم إضافية لقيمة أو سعر منزلها؟ كيف يمكنك الإجابة عن هذا السؤال؟

جدول رقم (6-2) انحدار أسعار البيع على عدد غرف النوم

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
القاطع (Intercept)	.28773.4327	4413.753	6.519	1.6E-10	20103.34	37443.53
X_2	13269.9801	1444.598	9.186	8.5E-19	10432.30	16107.66

يحتوي الانحدار البسيط هنا على بيانات عن سعر المنزل وعدد غرف النوم فقط. يمكنك أن تفكر بالأمر من خلال الاطلاع على بيانات كل المنازل في العينة وتوصلك إلى نتيجة مفادها أن المنازل التى بها غرف نوم أكثر تكون أكثر تكلفة . (على سبيل المثال، المنازل ذات الثلاث غرف نوم يكون سعرها أعلى من سعر نظيرتها ذات غرفتى النوم بما قيمته 13269.98 دولاراً).

وعلى الرغم من ذلك، فإن هذا لايعني بالضرورة أن زيادة غرفة نوم بالمنزل تؤدي إلى زيادة سعره بمقدار 13269.98 دولاراً. والسبب فى ذلك أن هناك العديد من العوامل الأخرى غير عدد غرف النوم لها إمكانية التأثير في أسعار المنازل. إضافة لذلك، هذه العوامل قد تكون ذات ارتباط عال ببعضها (بمعنى آخر، في الواقع أن المنازل الكبيرة يكون لها غرف نوم كثيرة، حمامات أكثر، أدوار أكثر ومساحة أكبر). وللتحقق من هذه الاحتمالية، دعنا نختبر أولاً مصفوفة الارتباط (انظر الفصل الثالث) لكل المتغيرات في هذا المثال (جدول رقم 6-3).

ولأن كل عناصر مصفوفة الارتباط موجبة، فإنه يتبع ذلك أن كل اثنين من المتغيرات ترتبط مع بعضها ارتباطا موجباً (على سبيل المثال، الارتباط بين عدد الحمامات وعدد غرف النوم هو 0.37، مما يشير إلى أن المنازل التى فيها حمامات أكثر يكون فيها غرف نوم أكثر). وفى مثل هذه الحالات، فإن الانحدار البسيط لايمكنه تحديد آثار المتغيرات المفردة على أسعار المنازل.

لذلك عندما تختبر طريقة الانحدار لكل المنازل وتأخذ فى الاعتبار أن المنازل التى فيها غرف نوم أكثر تكلف أكثر، وهذا لايعنى بالضرورة أن غرف النوم تضيف قيمة للمنزل. المشترون قد يكونون حقيقة يعطون قيمة للحمامات أو مساحة الأرض أكثر من غرف النوم. بمعنى آخر، المنازل التي فيها حمامات

أكثر قد تكون ذات قيمة أكثر. وتبقى الحقيقة، المنازل التي فيها حمامات أكثر لها أيضاً غرف نوم أكثر.

الانحدار البسيط ينظر لسعر المنزل وعدد غرف النوم ويرى أن المنازل التي فيها حمامات أكثر لها أيضاً غرف نوم أكثر. وبالتالي فهو يغفل حقيقة أن الأفراد يقيمون أو يعطون قيمة للحمامات. لذلك إذا نصحت صديقتك بأن إضافة غرفة نوم للمنزل سوف تضيف 13269.98 دولاراً فإنك بذلك تكون قد ضللتها بشكل كبير. فى الواقع، نقوم فى الانحدار البسيط، بحذف متغيرات تفسيرية هامة مثل مساحة الأرض، عدد غرف النوم، وعدد الأدوار. الانحدار يجمع بين إسهام كل هذه العناصر مع بعضها ويعطيها للمتغير المستقل الوحيد الذي تستطيع توزيعها له وهو: غرف النوم. وبالتالى فإن قيمة $\hat{\beta}$ كبيرة جدا [5].

في المقابل، يتيح لنا الانحدار المتعدد تحديد الإسهامات الفردية للمتغيرات التفسيرية الأربعة التي يفترض أن تؤثر في أسعار المنازل.

جدول رقم (6-3) مصفوفة ارتباط المتغيرات في مثال سعر المنازل

	سعر البيع	مساحة الأرض	عدد غرف النوم	عدد الحمامات	عدد الأدوار
سعر البيع	1				
مساحة الأرض	0.535795	1			
عدد غرف النوم	0.366447	0.151851	1		
عدد الحمامات	0.516719	0.193833	0.373768	1	
عدد الأدوار	0.41190	0.083674	0.407973	0.324065	1

الرقم $\hat{\beta}_2 = \$2,842.61$ يقترب إلى أن يكون مقياساً حقيقياً لأثر إضافة غرفة نوم، ومع ذلك، حتى هذا الانحدار المتعدد يبدو أنه يسقط بعض المتغيرات التفسيرية. وبإبراز هذا الرقم لصديقتك، فإنك سوف تكون واثقاً من أنك لاترتكب الخطأ المشار إليه أعلاه. أى أنه يمكنك أن تكون متأكداً أنه يبدو بشكل كبير أن غرفة النوم هى التى تضيف القيمة، وأنك لم توضح إسهامات المتغيرات التفسيرية المختلفة.

تحيز المتغيرات المحذوفة:

القضايا التي تمت مناقشتها في القسم السابق تتعلق بقضية إحصائية تسمى تحيز المتغيرات المحذوفة (Omitted Variables Bias). لن نقوم هنا بتطوير نظرية إحصائية ضرورية للتوضيح بشكل علمي ماذا يعني تحيز المتغيرات المحذوفة. وبشكل عام، فإنه يمكننا القول إذا حذفنا المتغيرات التفسيرية التي يجب إدراجها في الانحدار وإذا تم إجراء ارتباط بين هذه المتغيرات المحذوفة مع تلك المتغيرات المضمنة في الانحدار، فإن المعاملات للمتغيرات المضمنة سوف تكون خطأ. في المثال السابق، الانحدار البسيط لـ Y = سعر البيع على X = عدد غرف النوم، يستبعد العديد من المتغيرات التي كانت مهمة لتفسير أسعار المنازل (على سبيل المثال، مساحة الأرض، عدد غرف النوم،... الخ). هذه المتغيرات المحذوفة تم ربطها أيضاً أو إجراء ارتباط لها بعدد غرف النوم. لذا فإن تقدير المعامل $\hat{\beta} = 13,269.98$ لا يعتمد عليه نظراً لتحيز المتغيرات المحذوفة. الإجابة عن التساؤل حول لماذا يسبب حذف متغيرات هذا التحيز. على سبيل المثال، تُعد مساحة الأرض متغيراً تفسيرياً مهماً لأسعار المنازل، ولذلك "يجب" أن يتم إدراجها في الانحدار. إذا قمنا بحذفها من الانحدار، فإنها سوف تؤثر من خلال المدخل الوحيد الذي تستطيع الدخول من خلاله. وهو الارتباط الموجب بالمتغير

التفسيري الآخر: عدد غرف النوم. بمعنى آخر، معامل عدد غرف النوم سوف يكشف أثر غرف النوم ومساحة الأرض على أسعار المنازل.

هنالك أثر عملي آخر يتبع لتحيز المتغيرات المحذوفة يتمثل في أنه يجب عليك دائماً محاولة تضمين كل المتغيرات التفسيرية التي يمكن أن تؤثر على المتغير التابع. لسوء الحظ هذا نادراً ما يكون ممكنا في الواقع. أسعار المنازل، على سبيل المثال، تعتمد على العديد من المتغيرات التفسيرية الأخرى غير تلك الموجودة في ملف البيانات، HPRICE.xls (على سبيل المثال، وضع صيانة المنزل، لأي مدى يكون الجيران مريحين، توافر المساحة التخزينية، ما إذا كانت أرضيات المنزل من الخشب الصلب، نوعية الحديقة ... الخ).

في الواقع هناك العديد من المتغيرات التى تجمع عنها بيانات، ومعظمها يثار التساؤل حولها (على سبيل المثال، كيف يمكنك قياس مدى كون الجيران مريحين؟). سوف تقوم عملياً بحذف متغيرات وهناك القليل الذى يمكن عمله حيال ذلك، عدا بأن نأمل أن لايكون للمتغيرات المحذوفة قوة تفسيرية كبيرة وأنها غير مرتبطة بالمتغيرات التفسيرية المضمنة في الانحدار.

الفقرات السابقة تقدم تبريراً للتعامل مع أكبر عدد ممكن من المتغيرات التفسيرية. ومع ذلك، فهناك فكرة مغايرة لذلك تتمثل في استخدام أقل عدد ممكن من المتغيرات التفسيرية. ومن الواضح أن استخدام أو إدراج متغيرات غير مناسبة يقلل من دقة تقدير جميع المعاملات (حتى تلك التى تكون غير مناسبة). هذا الانخفاض في الدقة سوف ينعكس في القيم الكبيرة جداً لفترات الثقة والقيمة الاحتمالية.

كيف يمكننا الموازنة بين فوائد إدراج العديد من المتغيرات (على سبيل المثال، تخفيض المخاطرة الناجمة عن تحيز المتغيرات المحذوفة) و تكاليف محتملة لإدراج متغيرات غير مناسبة (بمعنى آخر، تخفيض دقة التقدير)؟ الممارسة الشائعة هى البدء

بأكبر عدد ممكن من المتغيرات التفسيرية، ثم التخلص من تلك المتغيرات غير المعنوية إحصائياً. وعلى الرغم من ذلك، إذا قمت بوضع عدد أكبر من المتغيرات التفسيرية لتبدأ بها، فإنك سوف تجد فعلياً أن كل المتغيرات التفسيرية غير ذات معنوية. لذلك، فإن الأمر يتطلب بعض الفهم البديهي حول ماهو أفضل انحدار مبدئي. المعنوية الإحصائية للمتغير التفسيري يمكن بالطبع تقييمها باستخدام القيم الاحتمالية التي تنتجها برامج الحاسب الآلي مثل إكسل. وبمجرد أن تتمكن من التخلص من المتغيرات التفسيرية غير المعنوية، فإنه يمكنك إجراء انحدار جديد يتضمن متغيرات تفسيرية أقل، وتكون قد تلافيت بشكل كبير مشكلة احتمال إدراج متغيرات ليس لها علاقة في الانحدار.

تمرين رقم (6-2):

استخدم مجموعة البيانات في الملف HPRICE.xls، واجعل Y = سعر المنزل لتكون المتغير التفسيري وخذ في الاعتبار المتغيرات التفسيرية المحتملة التالية:

X_1 = مساحة الأرض (بالأقدام المربعة).

X_2 = عدد غرف النوم.

X_3 = عدد الحمامات.

X_4 = عدد الأدوار (ماعدا القبو).

أ- قم بإجراء انحدار لـ Y على كل من X_1, X_2, X_3, X_4 (بمعنى آخر، قم بإعادة المثال أعلاه) وناقش النتائج التى توصلت إليها.

ب- قم بإجراء انحدار لـ Y على مختلف المجموعات الفرعية لكل من X_1, X_2, X_3, X_4 وناقش النتائج التى توصلت إليها.

ت- قم بمقارنة النتائج التي توصلت إليها في كل من أ و ب أعلاه، واختبر أثر حذف المتغيرات التفسيرية.

الارتباط الخطي:

الارتباط الخطى (Multicollinearity) هو مسألة إحصائية تتعلق بما سبق من نقاش. ويُعد الارتباط الخطي مشكلة تبرز إذا كان هناك ارتباط كبير بين جميع المتغيرات التفسيرية مع بعضها. وإذا كان هذا الوضع قائماً، تبرز صعوبة في فصل أثر المتغيرات التفسيرية على المتغير التابع. تكشف مشكلة الارتباط الخطى عن نفسها من خلال إحصائية t منخفضة ومن ثم قيم احتمالية عالية. في هذه الحالات، يمكنك القول بأن المعاملات ليست معنوية ولذلك يجب إسقاطها من الانحدار. في حالات نادرة، يمكنك أن تجد أن كل المعاملات غير معنوية باستخدام إحصائية t، بينما تكون قيمة R^2 كبيرة جداً وذات معنوية. بديهياً، ذلك يعنى أن المتغيرات التفسيرية مجتمعة تقدم الجزء الأكبر من القوة التفسيرية، إلا أن الارتباط الخطي يجعل من غير الممكن أن يحدد نموذج الانحدار ماهو المتغير التفسيري أو المتغيرات التفسيرية التى تؤثر فعلاً في المتغيرات التابعة.

ليس هناك الكثير الذى يمكن عمله لتصحيح هذه المشكلة أكثر من حذف بعض المتغيرات ذات الارتباط الخطي العالى من الانحدار. ومع ذلك هناك العديد من الحالات التى لاتود أن تفعل فيها ذلك. على سبيل المثال، في مثالنا عن أسعار المنازل، إذا وجدنا أن هناك ارتباطاً عالياً بين عدد غرف النوم وعدد الحمامات، فإن الارتباط الخطى يصبح مشكلة. إلا أنك قد تتردد في حذف أحد هذه المتغيرات ما دام الإحساس البديهي يشير إلى أن كليهما يؤثران في أسعار المنازل بشكل كبير. المثال التالي يوضح حالة توجد فيها مشكلة ارتباط خطي وكيف يمكن تصحيحها من خلال حذف متغير تفسيري.

مثال: أثر سعر الفائدة على سعر الصرف:

افترض أنك تريد اختبار أثر سياسة سعر الفائدة على سعر الصرف. أحد الطرق لذلك هى اختيار سعر صرف (على سبيل المثال سعر صرف الجنيه الاسترليني/الدولار) كمتغير تابع وإجراء انحدار له على سعر الفائدة. إلا أن هناك العديد من أسعار الفائدة التى يمكن استخدامها كمتغيرات تفسيرية (على سبيل المثال، سعر البنك المركزي، سعر أذونات الخزانة، ... الخ). أسعار الفائدة هذه تتشابه جداً مع بعضها وسوف تكون ذات ارتباط عال مع بعضها. إذا قمت بإدراج أكثر من سعر فائدة واحد منها فإنك قد تواجه مشكلة ارتباط خطى. طالما أن مختلف أسعار الفائدة هي بالضرورة تقيس نفس الظاهرة، الإحساس البديهي يشير إلى أن حذف كل متغيرات سعر الفائدة ماعدا واحد لن يؤدى إلى أى خسارة في القوة التفسيرية وسوف يحل مشكلة الارتباط الخطي. وعلى الرغم من ذلك، لن نعطي مثالاً رقمياً هنا ما دامت أسعار الفائدة وأسعار الصرف هي بيانات سلاسل زمنية. كما سوف نشاهد في الفصول اللاحقة، فإن الاستخدام غير السليم لأساليب الانحدار المتعدد مع بيانات السلاسل الزمنية قد ينتج عنه نتائج مضللة.

مثال: شرح الارتباط الخطي باستخدام بيانات افتراضية:

لتوضيح مشكلة الارتباط الخطي وكيف يمكن مواجهتها، سوف نقوم أولاً باستخراج 50 مشاهدة (N=50) بيانات افتراضية من نموذج الانحدار التالي:

$$Y = 0.5\, X_1 + 2\, X_2 + e$$

ونتوقع أن تكون تقديرات المربعات الصغرى الاعتيادية $\hat{\alpha} = 0$ ، $\hat{\beta}_1 = 0.5$ و $\hat{\beta}_2 = 2$ وهذه القيم قد تم استخدامها لإنشاء البيانات. وعلى الرغم من ذلك، فإن البيانات المستخرجة لديها ارتباط بين X_1، X_2 وهو ارتباط عال جداً. في الحقيقة أنه يساوي 0.98، مما يشير إلى أن الارتباط الخطى قد يكون مشكلة. جدول رقم (6-4) يعرض نتائج الانحدار باستخدام هذه البيانات.

جدول رقم (6-4) نتائج الانحدار باستخدام بيانات افتراضية

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
قاطع (intercept)	0.166191	0.1025278	1.57859	0.121137	-0.045601	0.377983
X_1	2.083733	0.952938	2.18664	0.033782	0.16667	4.00080
X_2	0.147775	0.965767	0.153013	0.879043	-1.7951	2.09065

$R^2 = 0.76$ وقيمة الاحتمال لاختبار أن $R^2 = 0$ هي 187-E-15.

جدول رقم (6-5) نتائج الانحدار باستخدام بيانات افتراضية تتجاهل X_2

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
قاطع (intercept)	0.166715	0.104146	1.60078	0.115989	-0.042685	0.376115
X_1	2.22690	0.178806	12.4543	1.2E-16	1.86739	2.58641

$R^2 = 0.76$ وقيمة الاحتمال في اختبار أن $R^2 = 0$ هي 1.2E-16

هذه النتائج مختلفة كثيراً عن التى كنا نأمل الحصول عليها. تقديرات المربعات الصغرى الاعتيادية تختلف عن تلك التى تم استخدامها لاستخراج البيانات. على سبيل المثال، $\hat{\beta}_1 = 2.08$ مع أن $\beta_1 = 0.5$ قد تم استخدامها

لاستخراج البيانات. فى الحقيقة، فإن تقدير المربعات الصغرى الاعتيادية $\hat{\beta}_1$ يكون تقريباً مساوياً للقيمة الحقيقية لـ β_2. هذه النتيجة توضح كيف تكون تقديرات المربعات الصغرى الاعتيادية "مشوشة" حول الدور الذى لعبته المتغيرات التفسيرية المفردة عندما تكون ذات ارتباط عال مع بعضها. لاحظ أيضاً أن أحد هذه المتغيرات التفسيرية ليس معنوياً إحصائياً عند مستوى معنوية 5% وأن الأخرى ذات معنوية هامشية فقط.

إضافة لذلك فإن 95% فترة ثقة لكل المعاملات تعد كبيرة جداً. هذه النتائج تشير إلى أن المتغيرات التفسيرية لديها قوة تفسيرية ضعيفة. وفى المقابل، فإن R^2 تكون كبيرة جداً وذات معنوية إحصائية قوية، وهى بذلك تشير إلى أن المتغيرات التفسيرية لديها قوة تفسيرية ممتازة.

وبالأخذ في الاعتبار مشكلة الارتباط الخطي، يؤيد العديد من الاقتصاديين القياسيين حذف X_2 من الانحدار. فإذا ما اتبعنا نصيحتهم ورجعنا للانحدار، فإننا سوف نحصل على النتائج الموضحة في جدول رقم (6-5).

لاحظ أن هذه النتائج تبدو جيدة إلى حد كبير من وجهة النظر الإحصائية. β_1 ذات معنوية إحصائية قوية وأن فترة الثقة تشير إلى أنه قد تم تقديرها بشكل دقيق. لذا فإن ذلك يعني أن حذف X_2 قد حل مشكلة الارتباط الخطى. المشكلة الوحيدة هى أن $\hat{\beta}_1$ لم تعد قريبة للقيمة الحقيقية 0.5 (وأن فترة الثقة لا تشتمل على 0.5). عموماً، طالما أن X_2 قد تم حذفها من النموذج، فإن X_1 تحاول أن تأخذ أثرها على المتغير التابع. ولأن X_1 ذات ارتباط عال جداً مع X_2، فإن X_1 قد تنوب عن X_2 بشكل كبير. لأن $\hat{\beta}_1$ تجمع بين تأثير كلا المتغيرين التفسيريين. بمعنى آخر، حذف متغيرات تفسيرية مهمة في

مثال أسعار المنازل أعطانا نتيجة متحيزةً لأثر غرف النوم على أسعار المنازل، حذف X_2 هنا أعطانا نتيجة متحيزةً لأثر X_1 على Y. ليس هناك حقيقة ما يمكنك عمله حيال ذلك أكثر من أن تأخذ في الاعتبار أن ذلك قد يحدث إذا كان هناك ارتباط خطى، وقم بتفسير النتائج التى توصلت إليها بحذر.

لاحظ أن الارتباط الخطي يتضمن ارتباطات بين المتغيرات التفسيرية، وليس المتغير التابع. ولكى تصبح هناك مشكلة، فإنه يجب أن تكون الارتباطات بين المتغيرات عالية جداً. إذا رجعنا لمثال تسعير المنازل، فإنه يمكننا أن نرى أن المتغيرات التفسيرية ذات ارتباط متوسط مع بعضها (على سبيل المثال، بعض الارتباطات هى حول 0.3 أو 0.4). إلا أن هذا الارتباط المتوسط لايؤدي إلى مشكلة ارتباط خطي ما دامت كل المعاملات تختلف بشكل كبير عن الصفر. (انظر القيم الاحتمالية في جدول رقم 6-5).

تمرين رقم (6-3):

لهذا التمرين استخدم مجموعة البيانات FOREST.xls التى فيها Y = الانحسار في مساحة الغابات، X_1 =الكثافة السكانية، X_2 = التغير النسبي في الأراضي الزراعية و X_3 = التغير النسبي فى أراضى المراعي. قم بإجراء تحليل انحدار متعدد لهذه المجموعة من البيانات تتناول القضايا أو المسائل التى تم طرحها في هذا الفصل. على سبيل المثال، قد تود أن تقوم بما يلي:

أ- إجراء انحدار لـ Y على X_1, X_2, X_3 وأن تشرح تقديرات المعاملات التى توصلت إليها.

ب- مناقشة المعنوية الإحصائية للمعاملات. هل هناك متغيرات تفسيرية يمكن حذفها؟

ج- مناقشة مدى الملاءمة في الانحدار.

د- حساب مصفوفة الارتباط. من خلال أخذ مدى الملاءمة ونتائج الانحدار في الاعتبار، مناقشة قضية الارتباط الخطى.

مثال: تكلفة إنتاج الكهرباء (تتمة ما ورد في الصفحات السابقة)

تعد المقدرة على شرح وتفسير نتائج الانحدار المتعدد من أهم المهارات التى يمكن أن يطورها الاقتصاديون التطبيقيون. فيما يلي نقدم مثالاً آخر، مع نتائج مكتوبة بالشكل الذى تكون عليه في تقرير مختصر.

تشير النظرية الاقتصادية الجزئية إلى أن تكلفة إنتاج المنشأة تعتمد على أسعار المدخلات (عناصر الإنتاج) المستخدمة في العملية الإنتاجية إضافة إلى حجم الإنتاج. لذلك، فإنه عند دراسة تكاليف الإنتاج، يجب أن تكون تكاليف الإنتاج هي المتغير التابع، والإنتاج وأسعار المدخلات يجب أن تكون هي المتغيرات التفسيرية. نستخدم هذه البيانات عن المتغيرات المذكورة لـ 123 منشأة كهربائية في الولايات المتحدة الأمريكية في العام 1970م (هذه البيانات متوافرة فى ملف ELECTRIC.xls). وتقيس هذه البيانات ما يلي:

Y= تكاليف الإنتاج (مقاسة بملايين الدولارات في السنة).

X_1 = حجم الإنتاج (مقاساً بآلاف الكيلووات ساعة (KWH) في السنة).

X_2 = تكلفة العمل (مقاسة بالدولارات لكل عامل في السنة).

X_3 = تكلفة رأس المال (مقاسة بالدولارات لكل وحدة من رأس المال في السنة).

X_4 = تكلفة الوقود (مقاسة بالدولارات لكل مليون BTU_s).

تتضمن نتائج الانحدار تلك المتغيرات الواردة في الجدول رقم (6-6). لاحظ أن كل المعاملات لديها الإشارة المتوقعة: زيادة الإنتاج أو أسعار أى من المدخلات سوف تزيد التكاليف. قيم المعاملات أيضاً مناسبة. وبذلك يمكن القول أن:

1- زيادة الإنتاج بمقدار ألف (KWH) سوف تزيد التكاليف بمقدار 4740 دولاراً، مع بقاء العوامل الأخرى على حالها. إننا واثقون بنسبة 95% بأن هذا الأثر الحدي هو على الأقل 4514 دولاراً وعلى الأكثر 4948 دولاراً.

2- زيادة الراتب السنوى للعامل المتوسط بمقدار دولار واحد سوف تزيد التكاليف بمقدار 3630 دولاراً في السنة، مع بقاء العوامل الأخرى على حالها. إننا واثقون بنسبة 95% بأن هذا الأثر الحدي هو على الأقل 1537 دولاراً وعلى الأكثر 5717 دولاراً. هذا يشير إلى درجة متوسطة من عدم التأكد، على الرغم من حقيقة أن هذا المعامل ذو معنوية قوية (p-value<0.01).

3- زيادة تكلفة رأس المال بمقدار دولار واحد لكل وحدة سوف يزيد التكاليف بمقدار 280080 دولاراً في السنة، مع بقاء العوامل الأخرى على حالها. فترة الثقة 95% لهذا المعامل هى أيضاً واسعة جداً.

4- زيادة تكلفة الوقود بمقدار دولار واحد لكل مليون *BTUs* سوف تزيد التكاليف بمقدار 783460 دولاراً في السنة، مع بقاء العوامل الأخرى على حالها.

طالما أن $R^2 = 0.94$، فإن المتغيرات التفسيرية مع بعضها مسئولة عن 94% من التغير في التكاليف. هذا رقم عال جداً وذو معنوية إحصائية قوية. وحقيقة أننا نقوم بتوضيح المتغير التابع تشير بشكل واضح إلى أنه من غير المناسب أن يتم حذف أى متغيرات تفسيرية هامة. إذا نظرنا إلى المعاملات بشكل منفرد، فإننا سوف نلاحظ أن القيم الاحتمالية جميعها ذات معنوية إحصائية عند مستوى 5%.

جدول رقم (6-6) نتائج الانحدار باستخدام بيانات تكلفة الكهرباء

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
قاطع (intercept)	-70.49511	12.69501	-5.55298	1.76E-07	-95.6347	-45.3556
X_1	0.00474	0.00011	43.22597	3.41E-74	0.004514	0.004948
X_2	0.00363	0.00106	3.43660	0.000814	0.001537	0.005717
X_3	0.28008	0.12949	2.16301	0.032557	0.023663	0.536503
X_4	0.78346	0.16579	4.72566	6.39E-06	0.455154	1.11177

$R^2 = 0.94$ وقيمة الاحتمال لاختبار $R^2 = 0$ هي 9.73E73.

توضح مصفوفة الارتباط في الجدول رقم (6-7) أن المتغيرات التفسيرية ليست ذات ارتباط قوي مع بعضها. أعلى ارتباط هو ذلك الذى بين تكلفة العمل وتكلفة الوقود وهو 0.32 فقط. الارتباطات الأخرى تقل كثيراً عنه مما يشير إلى عدم وجود مشكلة ارتباط خطي.

الاستثناء المحتمل للنتائج ذات المعنوية الإحصائية القوية بصفة عامة يتضمن X_3، تكلفة رأس المال. المعامل عن هذا المتغير له فترة ثقة واسعة قليلاً وقيمة احتمالية لاختبار الفرض أى $\beta_3 = 0$ تفوق قليلاً 3%. هذا يعني أنه لايمكننا رفض الفرض $\beta_3 = 0$ عند مستوى ثقة معنوية 1%. فى الواقع، أنك لاتستخدم

مستوى معنوية 1% (5% هو الأكثر شيوعاً). ومع ذلك، فإنه من أجل التوضيح أكثر دعنا نفترض إنك قد استخدمت مستوى ثقة معنوية 1%، فى هذه الحالة يمكنك القول أن β_3 لم تكن ذات معنوية إحصائية وقد تحذف تكلفة رأس المال كمتغير تفسيري. إذا قمت بعمل ذلك، ثم بعد ذلك قمت بإعادة إجراء الانحدار مستبعداً X_3 فإنك سوف تحصل على النتائج الموضحة في الجدول رقم (8-6). (لاحظ أنه ما دمنا نستخدم مستوى ثقة معنوية 1%، فإن الجدول الآن لديه فترة ثقة 99%).

الغرض الأساسي من الجدول رقم (8-6) شرح طريقة إحصائية شائعة (بمعنى آخر، إسقاط المتغيرات التفسيرية التي لا تكون ذات معنوية وإعادة إجراء الانحدار). ولأن حذف X_3 لم يغير النتائج كثيراً، فإننا لن نقوم بإعادة النقاش المشار إليه أعلاه عن تفسير النتائج.

جدول رقم (6-7) مصفوفة الارتباط للمتغيرات في بيانات تكلفة الكهرباء

سعر الوقود (price-fuel)	سعر رأس المال (price-capital)	سعر العمل (price-labor)	الإنتاج (Output)	
			1	الإنتاج
		1	0.056399	سعر العمل
	1	-0.078686	0.021481	سعر رأس المال
1	0.155224	0.318349	0.053507	سعر الوقود

جدول رقم (6-8) نتائج الانحدار باستخدام بيانات تكلفة الكهرباء بعد حذف X_3

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
قاطع (intercept)	-49.75804	8.449311	-5.889900	3.68E-08	-71.8765	-27.6396
X_1	0.004736	0.000111	42.6218	6.4E-74	0.004445	0.005027
X_2	0.003313	0.0001061	3.12145	0.002259	0.000535	0.006091
X_4	0.851586	0.165266	5.15282	1.03E-06	0.418956	1.284216

$R^2 = 0.94$ وقيمة الاحتمال لاختبار أن $R^2 = 0$ هي 3.5E-73.

ملخص الفصل:

1-نموذج الانحدار المتعدد مشابه إلى حد بعيد لنموذج الانحدار البسيط. هذا الفصل ركز فقط على الاختلافات بين النموذجين.

2- تفسير معاملات الانحدار يخضع لشروط بقاء العوامل الأخرى على حالها. على سبيل المثال، تقيس β_j الأثر الحدي لـ X_j على Y، مع بقاء المتغيرات التفسيرية الأخرى ثابتة.

3- إذا تم حذف متغيرات تفسيرية من الانحدار فإن المعاملات المقدرة قد تكون مضللة، وهذه الحالة تعرف بتحيز المتغيرات المحذوفة. وتتعقد المشكلة إذا كانت المتغيرات المحذوفة ذات ارتباط قوي مع المتغيرات المدرجة في نموذج الانحدار.

4- إذا كانت المتغيرات التفسيرية ذات ارتباط عال مع بعضها، فإن تقديرات المعاملات والاختبارات الإحصائية قد تكون مضللة. هذه الحالة تسمى بمشكلة الارتباط الخطي.

ملحق رقم (6-1): التفسير الرياضي لمعاملات الانحدار:

القراء الذين لهم إلمام ببعض حساب التفاضل يمكنهم استخدام تلك المعرفة للتوصل إلى الفروقات بين الانحدار البسيط والانحدار المتعدد من منظور رياضي. في حالة نموذج الانحدار البسيط، يمكن استخدام أساسيات حساب التفاضل لاشتقاق العلاقة التالية:

$$\frac{dX}{dX} = \beta$$

أي أن معامل الانحدار، β ، يمكن تفسيره كمقياس لمقدار التغيرات في Y إذا تغيرت X بمقدار صغير، وهذا اشتقاق كلي (total derivative). في حالة نموذج الانحدار المتعدد يمكننا القول أن:

$$\frac{\partial Y}{\partial Xj} = \beta_j$$

بمعنى آخر، أن المعاملات عبارة عن اشتقاقات جزئية (partial derivative)، وليست اشتقاقات كلية. هذا الاشتقاق الجزئي يمكن تفسيره كتقدير لأثر التغير في X_j على Y، مع الأخذ في الاعتبار أن كل المتغيرات التفسيرية الأخرى في حالة ثبات.

ملاحظات ختامية:

1- نظرياً، يجب علينا وضع "i" على كل المتغيرات لتعريف كل مشاهدة. بمعنى آخر، كان يجب علينا كتابة $Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + ei$. على الرغم من أن إضافة العديد من الحروف التعريفية يعد مربكاً ويجعل من الصعوبة قراءة المعادلة. لذلك، فإننا هنا، وفي باقي الكتاب، سوف نقوم بحذف الـ"i" (أو "t" في بيانات السلاسل الزمنية) ما لم يكن من المهم تحديد المشاهدة المفردة.

2- القراء الذين لديهم إلمام بحساب التفاضل يجب أن يأخذوا في الاعتبار أنه يمكننا إيجاد تقديرات المربعات الصغرى الاعتيادية في نموذج الانحدار المتعدد بالطريقة العادية. أي أنه يمكننا أخذ الاشتقاقات الأولى لكل من $\alpha, \hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k$ على أساس أنها تساوي صفراً، وبعد ذلك يتم الحل.

3- مصفوفة الجبر مهمة للاشتقاقات النظرية أو الاثباتات التي تتضمن نموذج انحدار متعدد، لأن المعادلة ستكون معقدة جداً بدونها. مصفوفة الجبر ليست أحد الجوانب التي يناقشها هذا الكتاب، أما إذا كنت تقوم بدراسة إضافية في الاقتصاد القياسي فإنك سوف ترى أهمية استخدامها.

4- الطرق التي تم مناقشتها للمتغير المستقل الواحد في ملحق رقم (5-1) تنطبق أيضاً على حالة العديد من المتغيرات التفسيرية. أي أن كل معامل سوف يكون له إحصاء t يمكن مقارنتها مع القيمة الحرجة 1.96 إذا كان حجم العينة كبيراً. في الحالات التي يكون فيها هناك العديد من المتغيرات التفسيرية، قد ترغب في اختبار فروض معقدة تتضمن معاملات عديدة (على سبيل المثال: $H_0: \beta_1 + \beta_2 = \beta_3$) هذه الاختبارات أكثر صعوبة عند إجرائها مقارنة بتلك التي تمت تغطيتها هنا. ومع ذلك، يمكنك الرجوع لملحق رقم (11-1)، المتضمن بعض النقاش عن اختبار الفرض لمثل هذه الحالات.

5- إذا وجدت أن هذا التبرير مربكاً، فإنه يمكنك الرجوع إلى الفصل الخاص بالارتباط. فقد أخذنا فيه مثالاً على المتغيرات، تدخين السجائر، شرب الكحول وسرطان الرئة. حيث

أشرنا إلى أن الدراسات العلمية أشارت إلى أن التدخين هو الذى يسبب سرطان الرئة، ومع ذلك يميل المدخنون أيضا لشرب الكحول أكثر من غير المدخنين، لذلك، فإن الارتباط بين شرب الكحول وسرطان الرئة هو ارتباط موجب على الرغم من أن شرب الكحول لايسبب سرطان الرئة. هذا النوع من القضايا هو تماماً من النوع الذي نجده في هذا المثال. أي أن الانحدار البسيط الذي يتضمن متغيرات سرطان الرئة وشرب الكحول فقط قد يشير إلى أن هناك أثراً كبيراً لشرب الكحول على سرطان الرئة، على الرغم من أن شرب الكحول لا يسبب سرطان الرئة.

لماذا يحدث ذلك؟ لأننا قد استبعدنا أو حذفنا متغير التدخين الذي هو متغير تفسيري مهم لسرطان الرئة. هذا المتغير التفسيري المستبعد له ارتباط مع المتغير التفسيري الذي تم استخدامه في الانحدار البسيط (بمعنى آخر، شرب الكحول).

الفصل السابع

الانحدار مع وجود المتغيرات الصورية
(Regression with Dummy Variables)

استخدمت الفصول السابقة بيانات كمية لتوضيح أهمية المفاهيم الإحصائية. بيد أن كثيراً من البيانات التي يستخدمها الاقتصاديون هي بيانات نوعية (انظر الفصل الثاني لمعرفة الفرق بين البيانات النوعية والبيانات الكمية). وتعد المتغيرات الصورية (الوهمية) (Dummy Variables) التي سبق شرحها بإيجاز في الفصل الثاني، وسيلة لتحويل المتغيرات النوعية إلى متغيرات كمية. وبمجرد أن تصبح البيانات كمية، يصبح من الممكن استخدام أساليب الارتباط والانحدار الموضحة في الفصول السابقة. اصطلاحاً، يعد المتغير الصوري متغيراً يأخذ إحدى القيمتين الصفر أو الواحد.

مثال: تفسير أسعار المنازل:

تناولنا في الفصل السابق مثالاً موسعاً لمعرفة العوامل التي تؤثر في أسعار المنازل في مدينة ويندسور في كندا. لاحظ أن المتغيرات التفسيرية التي استخدمناها في ذلك الفصل كانت جميعها كمية (مساحة الأرض بالأقدام المربعة، عدد الحمامات،... الخ). ولكن ثمة عوامل أخرى قد تؤثر في أسعار المنازل وإن كانت ليست كمية بصورة مباشرة. وتتضمن الأمثلة في هذا الصدد توافر: طريق إلى المبنى، مكيفات هواء، قاعة ترفيه، قبو وشبكة تدفئة مركزية تعمل بالغاز. جميع هذه المتغيرات عبارة عن متغيرات نوعية تأخذ صيغة نعم/لا (نعم= وجود طريق إلى المبنى، لا= عدم وجود طريق إلى المبنى).

لإجراء تحليل انحدار باستخدام هذه المتغيرات التفسيرية، سنقوم أولاً بتحويلها إلى متغيرات صورية من خلال تغيير "نعم/لا" إلى "1/صفر". وباستخدام الحرف D للإشارة إلى المتغيرات الصورية، بالإمكان تعريف المتغيرات التالية:

$D1$ = 1 في حالة توافر طريق إلى المبنى (صفر في حالة عدم توافر طريق).

$D2$ = 1 في حالة توافر قاعة ترفيه في المبنى (صفر في حالة عدم توافر قاعة ترفيه).

$D3$ = 1 في حالة توافر قبو في المبنى (صفر في حالة عدم توافر قبو).

$D4$ = 1 في حالة توافر شبكة غاز للتدفئة (صفر في حالة عدم توافر شبكة غاز للتدفئة).

$D5$ = 1 في حالة توافر أجهزة تكييف هواء (صفر في حالة عدم توافر أجهزة تكييف هواء).

فعلى سبيل المثال، ستكون قيم هذه المتغيرات فى مبنى له طريق وقبو وشبكة تدفئة مركزية تعمل بالغاز ولكن لايوجد أجهزة تكييف هواء ولا قاعة ترفيه:

$$D1 = 1, D2 = 0, D3 = 1, D4 = 1, D5 = 0.$$

توجد هذه المتغيرات في الملف HPRICE.XLS.

بمجرد تحويل المتغيرات النوعية إلى متغيرات صورية، يصبح بالإمكان تقدير الانحدار بالطريقة التقليدية وكذلك إمكانية استخدام كافة الأساليب النظرية والبديهيات التي تم مناقشتها في الفصول السابقة.

لماذا يتم تخصيص فصل بأكمله لهذا الموضوع؟ هناك إجابتان لهذا السؤال.

أولاً: يعد الانحدار مع استخدام متغيرات تفسيرية صورية شائعاً جداً وإن كان تفسير تقديرات المعاملات يختلف قليلاً. لذلك، يعد من الأفضل بحث كيفية التفسير بالتفصيل.

ثانياً: يعد الانحدار مع استخدام المتغيرات الصورية وثيق الصلة بمجموعة أساليب أخرى تسمى "تحليل التباين" (ANOVA). ومع أن هذه المجموعة من

الاساليب نادرة الاستخدام في علم الاقتصاد، إلا أنها عبارة عن أداة شائعة الاستخدام في العلوم الاجتماعية الأخرى والعلوم الطبيعية مثل علم الاجتماع، التربية، الإحصاءات الطبية وعلم الأوبئة.

وعلى الرغم من تميز معظم برامج الحاسب الآلي، مثل إكسل، بخصائص تحليل التباين، إلا أن مصطلح "تحليل التباين" يختلف عن ذلك المستخدم بواسطة الاقتصاديين ومن ثم فإن "تحليل التباين" قد يبدو غريباً وغير واضح بالنسبة لك (مثلاً، قائمة أدوات إكسل/تحليل البيانات بها خيارات ANOVA عديدة تشير إلى عامل واحد، عاملين مع التكرار، عاملين بدون تكرار). إن المطلوب ملاحظته هنا هو أن الانحدار باستخدام المتغيرات التفسيرية الصورية يستطيع القيام بكل ما يقوم به "تحليل التباين". وفي الواقع، يعتبر الانحدار مع متغيرات صورية أداة عامة وذات قوة أكبر مقارنة بتحليل التباين. فعلى سبيل المثال، فإن تحليل التباين أحادي العوامل أو ثنائي العوامل يشير إلى عدد المتغيرات التفسيرية الصورية. وليس بإمكان برنامج إكسل، (وكذلك معظم برامج الحاسب الآلي التي تحسب جدول تحليل التباين)، التعامل مع أكثر من عاملين في تحليل التباين. ورغم ذلك، فإن برنامج إكسل يتيح استخدام حتى 16 متغيراً تفسيرياً ضمن إمكانات البرنامج المتعلقة بالانحدار المتعدد، ومن ثم يمكنك من معالجة نماذج تحليل التباين الأكثر تعقيداً باستخدام أساليب الانحدار المتعدد مع متغيرات صورية. باختصار، إذا كنت تعرف كيفية استخدام أسلوب الانحدار وفهم نتائجه، فلا داعي أبداً لدراسة تحليل التباين.

تمرين رقم (7-1):

باستخدام ملف البيانات HPRICE.XLS، احسب وفسر الإحصاءات الوصفية ومصفوفة الارتباط للمتغيرات الصورية الخمسة الموضحة في المثال أعلاه. كيف يمكنك تفسير متوسط المتغير الصوري؟

الانحدار البسيط مع وجود متغير صوري:

سنبدأ بدراسة نموذج انحدار بسيط مع وجود متغير تفسيري صوري D:

$$Y = \alpha + \beta D + e$$

فإذا قمنا بتقدير المربعات الصغرى الاعتيادية لنموذج الانحدار أعلاه، سنحصل على قيمتي α و β. وبالإمكان النظر إلى فترات ثقة لكل من α و β ؛ اختبار قيم P لاختبار مدى المعنوية الإحصائية للمعاملات؛ حساب قيمة R^2 ؛ إجراء اختبار F لتحديد معنوية الانحدار؛.... الخ وذلك بنفس الطريقة التي استخدمت في الفصول السابقة (بإمكانك الرجوع إلى الفصول أرقام 4،5،6 إذا كنت لا تزال غير ملم بها). ومن الموضوعات المهمة في هذه المرحلة موضوع تفسير قيم المعاملات.

تعطى العلاقة الخطية المستقيمة بين Y و D القيمة الملائمة للمشاهدة رقم i في:

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} Di$$

ويجب ملاحظة أنه ونظراً لأنه قيمة D_i تكون صفراً أو واحداً، فإن $\hat{Y} = \hat{\alpha}$ أو $\hat{Y} = \hat{\alpha} + \hat{\beta}$. وسيكون من المهم هنا تقديم مثال لتوضيح كيفية استخدام هذه النتيجة في تفسير نتائج الانحدار.

مثال: تفسير أسعار المنازل

يوضح الجدول رقم (7-1)، نتائج لانحدار y= أسعار المنازل على D = المتغير الصوري لتكييف الهواء، باستخدام بيانات من الملف HPRICE.XLS.

لاحظ أن قيمة P أو فترة الثقة (الدنيا 95%، العليا 95%) توضح أن β ذات معنوية عالية. بالإضافة إلى ذلك، فإن $\alpha = 59.885$ و $\beta = 25.996$. كيف يمكننا شرح هذه الأرقام؟ بالطبع، يمكننا استخدام الأثر الحدي ذاته الذي استخدمناه في الفصل الرابع. أي يمكننا القول أن β هي قيمة مدى تغير Y عند تغير X بوحدة واحدة. ولكن مع المتغير الصوري الحالي فإن التغير بمقدار وحدة واحدة يعني ضمناً التغير من حالة "عدم توافر أجهزة تكييف هواء" إلى "توافر أجهزة تكييف". أي أنه يمكن القول أن المنازل التي بها أجهزة تكييف هواء تزيد قيمتها بمقدار 25.996 دولاراً عن "المنازل الخالية من أجهزة التكييف".

جدول رقم (7-1): انحدار أسعار المنازل على المتغير الصوري لتكييف الهواء

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
القاطع(intercept)	59884.85	1233.50	48.55	7.1E-200	57461.84	62307.8
D	25995.74	2191.36	11.86	4.9E-29	21691.18	30300.32

ومن ناحية أخرى، هناك وسيلة أخرى وثيقة الصلة للتفكير في نتائج هذا الانحدار عندما يكون المتغير التفسيري صورياً. ففي حالة المنازل الخالية من أجهزة التكييف، تكون قيمة D_i = صفراً وبالتالي فإن $\hat{Y} = 59.885$. بعبارة أخرى، يتوصل نموذج الانحدار إلى أن متوسط سعر المنازل الخالية من أجهزة التكييف يبلغ 59.885 دولاراً. وفى حالة المنازل المزودة بأجهزة تكييف،

$D_i = 1$ ؛ يتوصل نموذج الانحدار إلى أن: $\hat{Y}_i = \hat{\alpha} + \hat{\beta} = 85.881$. ومن ثم فـإن المنـازل المـزودة بـأجهزة تكييـف يبلـغ متوسـط سـعرها 85.881 دولاراً. وهذه إحدى الوسائل الأكثر شيوعاً لعرض نتائج الانحدار. وكبديل لذلك، يمكن التركيـز مباشـرة علـى $\hat{\beta}$ والقول أن أسـعار المنـازل المـزودة بـأجهزة تكييـف تكون أعلى بمقدار 25.996 دولاراً عن أسعار المنـازل التـي ليس فيهـا أجهـزة التكييف.

لمزيد من التوضيح، لاحظ أنه إذا لم نكن قد أجرينـا عمليـة الانحـدار ولكـن قمنـا بحسـاب متوسط أسـعار المنـازل المزودة بـأجهزة تكييـف، لوجدنا أن هـذه القيمة تساوي 85.881 دولاراً. وإذا كنا قد قمنا بحساب متوسط سعر المنـازل التـي لـيس فيهـا أجهـزة التكييـف، لوجدنا أنـه السـعر يسـاوي 59.885 دولاراً. وذلك يعنـي أننـا كنـا سنحصـل علـى النتـائج ذاتهـا التـي توصـلنا إليهـا مـن تحليـل الانحدار.

ورغم ذلك، يجب أن نتذكر النقاش السابق حـول تحيز المتغيرات المحذوفـة في الفصل السادس. إن مثال الانحدار البسيط هذا يحذف العديد مـن المتغيـرات التفسيرية المهمـة. وبالتأكيـد، لا يمكننـا اسـتخدام نتـائج هـذا الانحـدار البسيط لصياغة عبارات مثل "إن إضافة أجهزة تكييف إلـى منزلـك سـوف يزيـد سـعره بمقدار 25.996 دولاراً". ولمـا كانـت أجهـزة التكييـف لا تكلـف سـوى بضـع مئات أو آلاف من الدولارات، فإن مثل هذه العبارة تعد مضللة.

الانحدار المتعدد مع وجود متغيرات صورية:

سنقوم الآن بدراسة نموذج انحدار متعدد يحتوي على العديد من المتغيرات التفسيرية الصورية:

$$Y = \alpha + \beta_1 D_1 + ... + \beta_k D_k + e .$$

يمكن إجراء تقدير مربعات صغرى اعتيادية وتحليل إحصائي للنتائج بالطريقة التقليدية. ولتعزيز التفسير، سنعود مرة أخرى لمثال أسعار المنازل.

مثال: تفسير أسعار المنازل:

انظر في الحالة التي يكون فيها متغيران صوريان: $D_1 = 1$ في حالة أن يكون هناك طريق يؤدي إلى المنزل (D_1 = صفراً في حالة عدم توافر طريق)، $D_2 = 1$ في حالة توافر قاعة ترفيه في المنزل (D_2 = صفراً في حالة عدم توافر قاعة ترفيه). هذان المتغيران الصوريان يصنفان المنازل بوضوح في ملف البيانات إلى أربع فئات مختلفة:

1- منازل تتميز بتوافر طريق وقاعة ترفيه ($D_1 = 1$، $D_2 = 1$)

2- منازل تتميز بتوافر طريق ولكن لا توجد فيها قاعة ترفيه ($D_1 = 1$، D_2 = صفراً)

3- منازل لا يوجد طريق يؤدي إليها ولكن فيها قاعة ترفيه (D_1 = صفراً، $D_2 = 1$)

4- منازل لا طريق يؤدي إليها ولا توجد فيها قاعة ترفيه (D_1 = صفراً، D_2 = صفراً)

تذكر هذا التصنيف وأنت تقوم بتفسير الجدول رقم (7-2) الذي يعرض نتائج الانحدار الخاص بأسعار المنازل (Y) مع المتغيرين الصوريين D_1 و D_2.

وبتخصيص قيمة واحد أو صفر للمتغيرات الصورية، نحصل على قيم Y التوفيقية (المقدرة) في الفئات الأربعة من المنازل:

1- إذا كان $D_1 = 1$، $D_2 = 1$، فإن:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 = 47,099+21,160+16,024=84,283$$

بعبارة أخرى، يبلغ متوسط سعر المنزل الذي يتمتع بطريق يؤدي إليه وبقاعة للترفيه مبلغ 84.283 دولاراً.

2- إذا كان $D_1 = 1$ و $D_2 =$صفراً، فإن:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}_1 = 47,099 + 21,160 = 68,259 \; .$$

بعبارة أخرى، يبلغ متوسط سعر المنزل الذي يتمتع بطريق ولكن لا يحتوي على قاعة ترفيه مبلغ 68.259 دولاراً.

3- إذا كان $D_1 =$ صفر و $D_2 = 1$، فإن:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}_2 = 47,599 + 16,024 = 63,123$$

وهذا يعني أن متوسط سعر المنزل الذي لا يتمتع بطريق ولكن يحتوي على قاعة ترفيه يبلغ 63.123 دولاراً.

4- إذا كان $D_1 =$صفراً و $D_2 =$صفراً، فإن:

$$\hat{Y} = \hat{\alpha} = 47,099$$

أى أن متوسط سعر المنزل الذي لا يوجد طريق يؤدي إليه ولا توجد به قاعة ترفيه يبلغ 47.099 دولاراً.

خلاصة القول، يمكن استخدام الانحدار المتعدد مع وجود المتغيرات الصورية لتصنيف المنازل إلى فئات مختلفة وحساب متوسط السعر لكل فئة. وكبديل، يمكن عرض النتائج مباشرة من خلال تقديرات المعاملات، فعلى سبيل المثال، فإن β_1 عبارة عن مقياس للقيمة الحدية (الإضافية) لمنزل يتمتع بطريق بالمقارنة مع منزل لا يتمتع بمثل هذا الطريق مع ثبات المميزات الأخرى للمنزل (توافر قاعة ترفيه في المنزل أم لا).

جدول رقم (7-2): انحدار أسعار المنازل مع وجود المتغيرين الصوريين الطريق المؤدي للمنزل وقاعة الترفية.

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
قاطع (intercept)	47099.08	2837.62	16.60	2.42E-50	41525.02	52673.14
D_1	21159.91	3062.44	6.91	1.37E-11	15144.22	27175.60
D_2	16023.69	2788.63	5.75	1.52E-08	10545.86	21501.51

تمرين رقم (7-2):

اشرح المعلومات الإحصائية الواردة في الجدول رقم (7-2). هل جميع المتغيرات التفسيرية ذات معنوية إحصائية؟

تمرين رقم (7-3):

بالنسبة لهذا التمرين، استخدم Y= سعر المنزل، والمتغيرات الصورية $D_1 = 1$ في حالة توافر طريق يؤدي إلى المنزل ($D_1 =$صفراً إذا لم يكن هناك طريق) و $D_2 = 1$ إذا وجدت قاعة ترفيه بالمنزل ($D_2 =$صفراً إذا لم توجد قاعة ترفيه) وذلك من المثال الخاص بأسعار المنازل (يمكن الحصول عليها من الملف HPRICE.XLS). بدون استخدام أساليب الانحدار، احسب متوسط السعر للفئات الأربعة من المنازل الموضحة في المثال السابق.

ملاحظة: ما النتيجة المتحصل عليها إذا ضربت المتغير الصوري في Y؟ ماعلاقة هذه المتوسطات السعرية بمعاملات ونتائج الانحدار في المثال السابق؟

تمرين رقم (7-4):

بالنسبة لهذا التمرين، استخدم ملف البيانات HPRICE.XLS والمتغيرات الصورية الخمسة ($D_1,......, D_5$) الموضحة في بداية الفصل.

أ- كم عدد فئات المنازل الممكنة في ظل خمسة متغيرات صورية؟ (مثلاً: المنازل التي تتمتع بطرق مؤدية إليها، قاعة ترفيه، قبو، شبكة تدفئة مركزية تعمل بالغاز وأجهزة تكييف هواء، تمثل فئة واحدة). ما هي انعكاسات ذلك على تفسير نتائج الانحدار كما في المثال السابق؟

ب- كيف يمكنك حساب عدد المنازل في كل فئة باستخدام برنامج حاسب آلي مثل برنامج إكسل؟ مثلاً، من ضمن عدد المنازل البالغ 546 منزلاً في ملف البيانات، كم منزلاً يتمتع بطرق مؤدية إليه، وشبكة تدفئة مركزية تعمل بالغاز، وأجهزة تكييف هواء ولكن لا يوجد به قبو ولا قاعة ترفيه؟

ج- قم بإجراء الانحدار في Y= سعر المنزل بالنسبة للمتغيرات الصورية الخمسة.

د- ناقش المعنوية الإحصائية للمتغيرات التفسيرية.

هـ- احسب متوسط السعر لعدد قليل من أنواع المنازل المختارة (مثلاً، المنازل التي تتمتع بطرق مؤدية إليها و قاعة ترفيه وقبو ولكن لا توجد فيها شبكة تدفئة مركزية ولا أجهزة تكييف).

و- ماهى المزايا الأكثر أثراً في رفع سعر المنزل؟

الانحدار المتعدد مع وجود متغيرات تفسيرية صورية وغير صورية:

افترضنا في النقاش السابق أن كافة المتغيرات التفسيرية صورية. ولكن في الواقع نجد دائماً مزيجاً من الأنواع المختلفة للمتغيرات التفسيرية. ومن أبسط هذه الحالات، الحالة التي يكون فيها متغير صوري واحد (D) ومتغير تفسيري كمي واحد (X) في معادلة الانحدار:

$$Y = \alpha + \beta_1 D + \beta_2 X + e$$

بالإمكان تفسير نتائج هذه المعادلة في سياق المثال التالي:

مثال: تفسير أسعار المنازل:

إذا أجرينا معادلة الانحدار Y= سعر المنزل، على المتغير الصوري D = توافر جهاز تكييف، والمتغير التفسيري الكمي X= مساحة الأرض، فسوف نحصل على: $\alpha = 32693$، $\beta_1 = 20175$، و $\beta_2 = 5.638$. لقد لاحظنا أعلاه أن المتغير الصوري يأخذ فقط القيمة واحداً أو صفراً وأوضحنا أن قيمة "Y" المقدرة

قد تكون مختلفة لكل فئة من فئات المنازل. ومن ثم يمكن تفسير نتائج الانحدار على أنها توضح متوسط السعر للمنزل في كل فئة محتملة من تلك الفئات.

إن الأمور هنا قد لاتكون بالسهولة التي نتصورها ما دمنا نحصل على:

$$\hat{Y}_i = 52,868 + 5.638\, X_i$$

في حالة أن يكون D_i =1 (أي أن المنزل رقم i يحتوى على أجهزة تكييف هواء)

$$\hat{Y}_i = 32,693 + 5.638\, X_i \quad \text{و}$$

في حالة D_i = صفراً (لا يوجد مكيف هواء في المنزل).

بعبارة أخرى، هناك خطا انحدار مختلفان استناداً إلى توافر أو عدم توافر مكيف هواء في المنزل. قارن هذا مع النقاش الوارد في المثال السابق الذي يتضمن متغيراً تفسيرياً صورياً واحداً فقط. فقد أوضح الانحدار في تلك الحالة أن متوسط سعر المنزل اختلف بالنسبة للمنزل الذي فيه مكيف هواء عنه في المنزل الذي ليس فيه مكيف هواء. هنا يمكن القول أنه يوجد خط انحدار مختلف تماماً. بعبارة أخرى لايمكننا بسهولة تحديد متوسط سعر المنزل (كما فعلنا في الأمثلة السابقة في هذا الفصل) لمختلف فئات المنازل.

ورغم ذلك، يمكننا القول بأن $\beta_1 = 20,175$ يمثل مقياساً للقيمة الحدية (الإضافية) في قيمة المنزل نتيجة تزويده بمكيف هواء مع ثبات بقية المزايا. بعبارة أخرى، إذا قارنا منزلين متماثلين من حيث قيمة X (مساحة الأرض)، فإن قيمة $\hat{Y}_i$ ستكون دائماً أكبر بمبلغ 20.175 دولاراً بالنسبة للمنزل المزود بمكيف هواء مقارنة بالمنزل الذي لا يتوافر فيه مكيف هواء.

من الأهمية بمكان إجراء اختبار أكثر دقة لخطي الانحدار في الحالتين. لاحظ أن كليهما له درجة الميل ذاتها $(\hat{\beta}_2 = 5.638)$ ويختلفان فقط في حجم القاطع (إذا كان D_i =1، فإن القاطع يكون 52.868. أما إذا كان D_i = صفراً، فإنها تكون 32.693). ونظراً لتساوي الميل في الخطين (الميل هو التأثير الحدي)، فإن التأثير الحدي لمساحة الأرض على سعر المنزل تكون مماثلة للمنازل المزودة وغير المزودة بأجهزة تكييف. فعلى سبيل المثال يمكن القول بأن "زيادة مقدارها قدم مربع واحد في مساحة الأرض ستكون مصحوبة بزيادة مقدارها 5.63 دولاراً في سعر المنزل".

بالإمكان تطبيق النقاش السابق على حالة تعدد المتغيرات التفسيرية الصورية وغير الصورية. يتضمن المثال التالي نموذج معادلة انحدار بمتغيرين تفسيريين صوريين ومتغيرين تفسيريين غير صوريين:

$$Y = \alpha + \beta_1 D_1 + \beta_2 D_2 + \beta_3 X_1 + \beta_4 X_2 + e.$$

يتضمن تفسير النتائج لهذا النموذج عناصر من كافة الأمثلة السابقة فى هذا الفصل.

مثال: تفسير أسعار المنازل: (تابع للمثال أعلاه)

إذا كان Y= سعر المنزل.

D_1 = المتغير الصوري لتوافر طريق يؤدى للمنزل.

D_2 = المتغير الصوري لتوافر قاعة ترفيه في المنزل.

X= مساحة الأرض.

X_2 = عدد غرف النوم.

فسوف نحصل على ما يلي:

$$\hat{B}_4 = 10{,}562 ، \hat{\beta}_3 = 5.197 ، \hat{\beta}_2 = 10{,}969 ، \hat{\beta}_1 = 12{,}598 ، \hat{\alpha} = -2{,}736$$

بإمكاننا تفسير هذه النتائج من خلال تحديد معادلات الانحدار المقدرة $(\hat{Y})$ لمختلف القيم الممكنة للمتغيرات الصورية.

1- إذا كان $D_1 = 1$ و $D_2 = 1$ فإن:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 X_1 + \hat{\beta}_4 X_2 = 20{,}831 + 5.197 X_1 + 10{,}562 X_2$$

وهذا هو خط الانحدار الخاص بمنزل له طريق وبه قاعة ترفيه.

2- إذا كان $D_1 = 1$ و $D_2 =$ صفراً فإن:

$$\hat{Y} = 9{,}862 + 5.197 X_1 + 10{,}562 X_2$$

وهذا هو خط الانحدار الخاص بمنزل له طريق ولكن لاتوجد به قاعة ترفيه.

3- إذا كان $D_1 =$ صفراً و $D_2 = 1$ فإن:

$$\hat{Y} = 8{,}233 + 5.197 X_1 + 10{,}562 X_2$$

وهذا هو خط الانحدار الخاص بمنزل به قاعة ترفيه ولا يوجد طريق يؤدي إليه.

4- إذا كان $D_1 =$ صفراً و $D_2 =$ صفراً فإن:

$$\hat{Y} = -2{,}736 + 5.197 X_1 + 10{,}562 X_2$$

وهذا هو خط الانحدار الخاص بمنزل لا يوجد فيه قاعة ترفيه ولا يوجد طريق يؤدي إليه.

وهذا يعني أنه يمكن استخدام متغيرين صوريين للحصول على أربعة خطوط انحدار مختلفة. وهذه الخطوط متساوية من حيث الميل ولكن مختلفة من حيث نقاط التقاطع. تقيس المعاملات للمتغيرات الصورية $(\hat{\beta}_1, \hat{\beta}_2)$ القيمة الإضافية لتوافر طريق يؤدي للمنزل وتوافر قاعة ترفيه فيه. وتعتبر معاملات المتغيرات غير الصورية $(\hat{\beta}_3, \hat{\beta}_4)$ بمثابة الأثر الحدي لمساحة الأرض وعدد غرف النوم.

وفيما يلي عدد من العبارات التي يمكن صياغتها بخصوص نتائج الانحدار:

1- المنازل التي يتوافر لها طرق مؤدية إليها تكون قيمتها أعلى بمقدار 12.598 دولاراً بالمقارنة مع منازل مماثلة ولكن لاتوجد طرق مؤدية إليها.

2- إذا أخذنا في الاعتبار منازل متساوية من حيث عدد غرف النوم، فإن إضافة قدم مربع واحد لمساحة الأرض سوف تزيد سعر المنزل بمقدار 520 دولاراً.

3- إن إضافة غرفة نوم واحدة للمنزل سوف تزيد سعره بمقدار 10.562 دولاراً مع بقاء العوامل الأخرى ثابتة.

ورغم ذلك يجب التأكيد على أن كافة هذه العبارات تفترض أن تحيز المتغيرات المحذوفة لا يعد مشكلة في الانحدار. علاوة على ذلك، فإن العبارة التي تحتوى ضمناً على السببية (مثل: إن إضافة قدم مربع واحد إلى مساحة الأرض سوف يزيد سعر المنزل بمقدار 520 دولاراً) تعد صحيحة فقط إذا كانت تمثل الحالة التي يكون فيها المتغير التفسيري سبباً لحدوث المتغير التابع (انظر الفصلين الرابع والسادس للمزيد من التوضيح عن السببية في الانحدار).

تمرين رقم (7-5):

بالنسبة لهذا السؤال استخدم ملف البيانات HPRICE.XLS، المتغيرات الصورية الخمسة $D_1,......,D_5$ الواردة في التمرين رقم (7-4) بالإضافة إلى المتغيرات التفسيرية غير الصورية:

X_1 = مساحة أرض المنزل بالأقدام المربعة.

X_2 = عدد غرف النوم.

X_3 = عدد الحمامات.

X_4 = عدد الأدوار (بدون القبو).

أ- احسب الانحدار Y للمتغيرات: $D_1,......,D_5, X_1,.....,X_4$

ب- حدد أي المتغيرات هو الأكثر أهمية من الناحية الإحصائية.

ج- أي الخصائص التي تم قياسها بواسطة المتغيرات الصورية هي الأكثر تأثيراً في أسعار المنازل؟

د- اختر تشكيلات محددة من المتغيرات الصورية (مثلاً تشكيلة تشير إلى منزل يتميز بـ: طريق وقبو وليس به تدفئة مركزية ولا يتوافر فيه أجهزة تكييف ولا قاعة ترفيه) واكتب المعادلة الخاصة بخط الانحدار.

هـ- ناقش النتائج المتعلقة بالمتغيرات التفسيرية غير الصورية مع الإشارة بصفة خاصة للأوضاع الأخرى التي تظل ثابتة (دون تغيير).

تفاعل المتغيرات الصورية وغير الصورية:

لقد استخدمنا المتغيرات الصورية أعلاه بطريقة تسمح للقاطع (α) بأخذ قيم مختلفة في خط الانحدار مع ثبات ميل خط الانحدار. ورغم ذلك، يمكن السماح بدرجات ميل مختلفة من خلال التفاعل بين المتغيرات الصورية والمتغيرات غير الصورية. ولفهم ذلك، انظر نموذج الانحدار التالي:

$$Y = \alpha + \beta_1 D + \beta_2 X + \beta_3 Z + e.$$

D و X هما المتغير التفسيري الصوري والمتغير التفسيري غير الصوري، كما ورد أعلاه. ولكننا أضفنا هنا متغيراً جديداً (Z) يمكن تعريفه كما يلي: Z= DX. كيف يمكننا تفسير نتائج معادلة الانحدار Y مع المتغيرات D,X,Z؟ بالإمكان الإجابة عن هذا السؤال من خلال الإشارة إلى أن قيمة Z إما صفر (بالنسبة للمشاهدات التى يكون فيها D = صفراً) أو X (للمشاهدات التى يكون فيها D = 1). فإذا نظرنا إلى خطوط الانحدار المقدرة مع D = صفراً و D =1، فإننا نحصل على الآتي:

إذا كان D = 1 فإن:

$$\hat{Y} = (\hat{\alpha} + \hat{\beta}_1) + (\hat{\beta}_2 + \hat{\beta}_3)X$$

وإذا كان D = صفراً فإن:

$$\hat{Y} = \hat{\alpha} + \beta X$$

بعبارة أخرى، هناك خطا انحدار مختلفان لكل من D = صفراً و D =1. وهذان الخطان لهما قاطعان مختلفان ودرجات ميل مختلفة. وهذا يعني أن التأثير الحدي لـ "X" على Y يختلف في حالة D = صفراً وحالة D =1.

وبالإمكان في التقارير كتابة كل من خطي الانحدار بصورة مستقلة باستخدام المصطلحات والتفسيرات الواردة في الفصلين الرابع والسادس.

مثال: تفسير أسعار المنازل:

إذا أجرينا معادلة الانحدار: Y= سعر المنزل باستخدام ثلاثة متغيرات تفسيرية: D = المتغير الصوري لأجهزة تكييف الهواء، X= مساحة الأرض، Z= DX، فإننا سنحصل على: $\alpha = 35.684$، $\hat{\beta}_1 = 7.613$، $\hat{\beta}_2 = 5.02$، $\hat{\beta}_3 = 2.25$. وهذا يعني ضمناً أن التأثير الحدي لمساحة الأرض على سعر المنزل يبلغ 7.27 (بمعنى أن إضافة قدم مربع واحد لمساحة المنزل ستكون مصحوبة بزيادة مقدارها 7.27 دولاراً في سعر المنزل) بالنسبة للمنازل المزودة بمكيفات هواء، وزيادة مقدارها 5.02 دولار فقط بالنسبة للمنازل التي ليس فيها أجهزة التكييف. علاوة على ذلك، ولما كانت قيمة P المقابلة لـ $\hat{\beta}_3$ تساوي 0.02 فإن الاختلاف في التأثير الحدي يعد معنوياً من الناحية الإحصائية. وهذه النتيجة تبين إلى أن زيادة مساحة الأرض تسهم بزيادة أكبر في سعر المنزل المزود بمكيف هواء مقارنة بسعر المنزل الذي ليس فيه مكيف هواء.

تمرين رقم (7-6):

بالنسبة لهذا التمرين، استخدم ملف البيانات HPRICE.XLS وخمسة متغيرات صورية $(D_1,......,D_5)$ وأربعة متغيرات غير صورية $(X_1,....,X_4)$ حسب الموضح في التمرين رقم (7-4). اختبر تشكيلات

مختلفة من هذه المتغيرات التفسيرية مع بعض الحدود التفاعلية بين المتغيرات الصورية وغير الصورية (مثلا: حاول إضافة 10 متغيرات تفسيرية: $X_1,....,X_4$ والمتغيرات الأربعة غير الصورية: $D_1,......,D_5$ زائداً حداً تفاعلياً، D_1X_2 ، مثلاً). هل بإمكانك الحصول على أى حدود تفاعلية (Z_s) ذات معنوية إحصائية؟ اشرح النتائج التي تحصل عليها.

تمرين رقم (7-7):

يحتوى الملف (WAGEDISC.XLS) على بيانات حول مجموعة موظفين تضم 100 موظف (N=100) في مهنة معينة. افترض أننا نريد تحديد العوامل التي توضح الفوارق في الرواتب في هذه المهنة مع العمل على معالجة قضية التمييز من حيث جنس الموظف (ذكر أو أنثى) في هذه المهنة. ضمن ملف البيانات المتغيرات التالية:

Y = الراتب (بآلاف الدولارات)

X_1 = المستوى التعليمي (عدد سنوات الدراسة)

X_2 = الخبرة العملية (سنوات العمل)

D = المتغير الصوري الخاص بالجنس (1= ذكر، صفر= أنثى)

أ- احسب الإحصاءات الوصفية لمجموعة البيانات هذه. مثلاً، ماهو متوسط الراتب؟

ب-احسب متوسط الراتب للموظفين الذكور والإناث، كل على حدة وقارن بينهما.

جـ- قم بإجراء عملية انحدار بسيط (Y) بالنسبة للمتغير D. هل معامل الميل في هذا الانحدار ذو معنوية إحصائية؟ قارن نتيجة الانحدار مع النتيجة التي حصلت عليها في (ب). هل يمكنك استخدام هذه النتائج لاستنتاج أن هناك تمييزاً ضد المرأة في هذه المهنة؟

د - قم بإجراء عملية انحدار متعدد (Y) بالنسبة للمتغيرات D, X_2, X_1 واكتب تقريراً مختصراً يوضح النتائج التي حصلت عليها ومعالجة قضية التمييز في الرواتب في هذه المهنة. هل تعد نتائجك ذات معنوية إحصائية؟

هـ- قارن نتائج الفقرة (د) مع نتائج الفقرة (جـ). لماذا تعد هذه النتائج مختلفة؟

ملاحظة: احسب مصفوفة ارتباط لكافة المتغيرات التفسيرية وفكر جيداً فى معنى هذه الارتباطات.

و- أوجد متغيراً جديداً (Z=DX) وأجر عملية انحدار (Y) للمتغيرات $Z \cdot D \cdot X_2 \cdot X_1$. هل يعد Z ذا معنوية إحصائية؟ ما هو التغير الذي سيطرأ على التقرير الذي أعددته في الفقرة (د)؟ وضح ما الذي يقيسه معامل Z؟

ماذا يعني إذا كان المتغير التابع متغيراً صورياً؟

فيما سبق كان تركيزنا منصباً على الحالة التي يمكن أن تكون فيها المتغيرات التفسيرية صورية. إلا أنه في بعض الحالات يمكن أن يكون المتغير التابع صورياً. فعلى سبيل المثال، قد يهتم خبير في اقتصاديات النقل بدراسة خيارات الأفراد بين النقل العام والسيارات الخاصة. وقد يتضمن التحليل العملي جمع

بيانات من العديد من الأفراد حول عاداتهم فيما يختص بالتنقل. وقد تتضمن المتغيرات التفسيرية المحتملة ما يلي: زمن الانتقال، دخل الفرد، ... الخ. ورغم ذلك، فإن المتغير التابع سيكون نوعياً (أي، يقول كل فرد: "نعم، أستخدم سيارتي في الذهاب إلى مكان العمل" أو "لا أستخدم سيارتي في الانتقال إلى مكان العمل"). وينبغي على الاقتصادي إنشاء "متغيراً تابعاً صورياً".

إن الأساليب المستخدمة مع المتغيرات الصورية التابعة تعد خارج نطاق هذا الكتاب[1]. ورغم ذلك، هناك أمران جديران بالاهتمام.

1- هناك بعض المشكلات المصاحبة لاستخدام طريقة المربعات الصغرى الاعتيادية فى هذه الحالة، ولكن هذه المشكلات ليست كثيرة، ومن ثم فإن تقدير المربعات الصغرى الاعتيادية قد يكون ملائماً في الكثير من الحالات.

2- على الرغم من ذلك، هناك أساليب تقدير أفضل من طريقة المربعات الصغرى الاعتيادية. يتمثل البديلان الرئيسان في Logit وProbit. إن برامج الحاسب الآلي الإحصائية المحدودة (مثل إكسل) لا تستطيع القيام بتنفيذ هذه الأساليب المتعلقة بالتقديرات. لذلك إذا أردت العمل بتوسع مع نماذج المتغيرات التابعة الصورية، فيجب عليك استخدام برامج مختلفة مثل LIMDEP (وهو الأفضل)، انظر الموقع http://www.Limdep.com لمزيد من التفاصيل. ويعد مثل هذا البرنامج صعب الاستخدام إلى حدٍ ما بالنسبة للمبتدئين.

ملخص الفصل:

1- تأخذ المتغيرات الصورية القيمة (واحد أو صفر) غالباً ما تستخدم مع البيانات النوعية.

2- إن الأساليب الإحصائية التي تستخدم مع المتغيرات التفسيرية الصورية مماثلة لتلك التي تستخدم مع المتغيرات التفسيرية غير الصورية.

3- إن معادلة الانحدار التي تتضمن متغيرات تفسيرية صورية فقط، تصنف المشاهدات بصورة واضحة إلى فئتين مختلفتين (المنازل التى فيها مكيفات هواء والمنازل التي ليس فيها مكيفات هواء). ويكون تفسير النتائج بشكل دقيق بماهية الفئات المستخدمة.

4- إن معادلة الانحدار التي تتضمن متغيرات تفسيرية صورية وغير صورية، تصنف المشاهدات بوضوح إلى فئات وتشير إلى أن كل فئة سيكون لها خط انحدار بنقطة تقاطع مختلفة ولكن جميع الخطوط متساوية من حيث درجة الميل.

5- إن معادلة الانحدار التي تحتوى على متغيرات تفسيرية صورية وغير صورية وتفاعلية (المتغيرات الصورية مضروبة في غير الصورية) تصنف المشاهدات بوضوح إلى فئات وتشير إلى أن كل فئة سيكون لها خط انحدار مختلف وذات قواطع ودرجات ميل مختلفة.

6- إذا كان المتغير التابع صورياً، فلابد من استخدام أساليب أخرى (لايناقشها هذا الكتاب).

ملاحظات ختامية:

1- تعرف مثل هذه النماذج بنماذج المتغير التابع المحدود (Limited Dependent Variables) بمعنى أن المتغير التابع يأخذ مدى محدوداً من القيم.

الفصل الثامن

الانحدار مع وجود فترات تباطؤ:
نماذج فترات التباطؤ الموزعة

(Regression with Time Lags: Distributed Lag Models)

تهتم العديد من فروع علم الاقتصاد (مثل علم الاقتصاد الكلي والتمويل) بتحليل بيانات السلاسل الزمنية. ولكن ربما قد لاحظت أن كافة الأمثلة الواردة في الفصول (3-7) تستخدم البيانات المقطعية. لقد تعمدنا حتى الآن عدم مناقشة بيانات السلاسل الزمنية (Time Series Data) لكونها تحتوي على قضايا لا تحدث مع حالة البيانات المقطعية. إن الهدف من هذا الفصل هو التمهيد لهذه القضايا وتوضيح سبب تخصيصنا لما يقارب نصف هذا الكتاب لموضوع السلاسل الزمنية بوصفها مجالاً مختلفاً عن الانحدار المتعدد. سنقوم بعد هذه المقدمة التمهيدية بدراسة أبسط أسلوب للتعامل مع السلاسل الزمنية: "نموذج فترات التباطؤ الموزعة" (Distributed Lag Model).

لا يختلف هدف الاقتصادي الذي يتعامل مع بيانات السلاسل الزمنية كثيراً عن هدف الاقتصادي الذي يتعامل مع البيانات المقطعية، فكلاهما يسعى لبناء وتقدير معادلة انحدار تربط متغيراً تابعاً ببعض المتغيرات التفسيرية. ولكن الاقتصادي الذي يستخدم بيانات سلسلة زمنية سوف يواجه مشكلات لا يواجهها الذي يستخدم البيانات المقطعية يمكن حصرها في: (1) إن متغير سلسلة زمنية قد يؤثر في متغير آخر بفترة تباطؤ، (2) إذا كانت المتغيرات غير مستقرة (nonstationary) ستبرز مشكلة "الانحدار الزائف" (Spurious Regression)

ليس من المتوقع في هذه المرحلة، أن تفهم المشكلة الثانية من هاتين المشكلتين. سيتم بحث المصطلحات: غير مستقرة، ومستقرة (Stationary) وانحدار زائف بحثاً مفصلاً في الفصول اللاحقة من هذا الكتاب. ولكن تذكر جيداً هذه القاعدة العامة: إذا كانت لديك متغيرات سلسلة زمنية غير مستقرة، فلا ينبغي عليك إدراجها في نموذج الانحدار. فالحل الأفضل هو تحويل هذه المتغيرات إلى متغيرات مستقرة قبل إجراء عملية الانحدار. ولكن هناك استثناء واحد من هذه القاعدة العامة (سوف نناقشه لاحقاً) ويتم عادة في حالة أن تكون المتغيرات

المدرجة في نموذج الانحدار متكاملة (Cointegrated). وسوف نتناول بإسهاب، لاحقاً، ما نعنيه بهذه المصطلحات. فإذا كنت تعتقد أن تناولها حالياً بدون تعريفات قد تربك فهمك، فقط فكر بالأمر حسب المنظور التالي: تظهر بعض المشكلات مع بيانات السلاسل الزمنية ولكنها لا تحدث مع البيانات المقطعية. وهذه المشكلات تجعل من الصعوبة بمكان توظيف الانحدار المتعدد على الصيغة الواردة في الفصول (4-7). إن الهدف من الفصول الأربعة اللاحقة هو استعراض كيفية الاستخدام الصحيح للانحدار المتعدد مع بيانات السلاسل الزمنية.

سنفترض في هذا الفصل أن كافة المتغيرات في الانحدار مستقرة، وسوف نوضح في الفصل التالي ماذا يعني ذلك. لاحظ فقط عند هذه النقطة أن المشكلة الثانية لن تحدث ومن ثم يمكننا التركيز على المشكلة الأولى فقط.

يمكن فهم المشكلة الأولى بوضوح من خلال تقديم بعض الأمثلة البسيطة. عند تقديرنا لنموذج انحدار، يتركز اهتمامنا على قياس تأثير متغير تفسيري واحد أو أكثر على المتغير التابع. ولكن ينبغي أن نكون حذرين عند اختيار المتغيرات التفسيرية ما دام تأثيرها في المتغير التابع قد لا يظهر إلا بعد فترة من الزمن.

فعلى سبيل المثال، إذا توقع البنك المركزي حدوث ركود في الأنشطة الاقتصادية، فقد يلجأ إلى تخفيض سعر الفائدة. ولكن تأثير تخفيض سعر الفائدة لن يظهر على النشاط الاقتصادى إلا بعد عام كامل أو أكثر وكذلك على المتغيرات الأخرى المهمة (مثل معدل البطالة). وبصفة عامة، فإن كافة أدوات السياسات النقدية والمالية لا تظهر آثارها إلا في المستقبل. وهذه المشكلة تظهر أكثر في مجال الاقتصاد الكلى وإن كان من المحتمل حدوثها أيضاً في مجال الاقتصاد الجزئي. ويمكن أخذ المثال التالي في هذا الصدد: إن اتخاذ شركة ما قراراً بضخ استثمارات جديدة (شراء أجهزة حاسب آلي مثلاً) لن يؤثر فوراً في الإنتاج. فشراء الحاسبات وتركيبها وتدريب الموظفين على استخدامها سوف

يتطلب بعض الوقت. إن تأثير مثل هذا الاستثمار في الإنتاج لن يظهر إلا في فترة زمنية مستقبلية.

لترجمة هذا المفهوم إلى لغة الانحدار، ينبغي أن تعتمد قيمة المتغير التابع عند نقطة زمنية محددة ليس على قيمة المتغير التفسيري عند تلك النقطة فحسب، بل أيضاً على قيم المتغيرات التفسيرية في الماضي. يعرف النموذج الأبسط لاستيعاب هذه التأثيرات بنموذج فترات التباطؤ الموزعة (Distributed Lag Model). وهو عبارة عن نموذج انحدار يأخذ الصيغة التالية: [1]

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + ... + \beta_q X_{t-q} + e_t.$$

وهذا النموذج مماثل تماماً لنموذج الانحدار المتعدد الذي تم مناقشته في الفصل السادس باستثناء أن "المتغيرات التفسيرية" ليست مختلفة كلياً (مثلاً، مساحة الأرض، عدد الحمامات، عدد غرف النوم،... الخ) ولكنها مجرد متغير تفسيري واحد تم مشاهدته في فترات زمنية مختلفة. تشير المتغيرات في الطرف الأيمن من هذا النموذج إلى متغيرات متباطئة (Lagged Variables) ويمثل "q" ترتيب أو درجة المتباطئة (Lag Order) أو طول المتباطئة (Lag Length). سنركز على الحالة التي يعتمد فيها المتغير التابع على متغير تفسيري واحد والمتباطئات الخاصة به. ورغم ذلك، يمكن تعميم ذلك على متغيرات تفسيرية عديدة جميعها ذوات فترات تباطؤ.

وبما أن نموذج فترات التباطؤ الموزعة عبارة عن نموذج انحدار، فإن كل ما ذكرناه في الفصول (4-6) بشأن الانحدار يعد وثيق الصلة بذلك. فعلى سبيل المثال، بإمكان برامج الحاسب الآلي المتخصصة، مثل إكسل، حساب تقديرات المربعات الصغرى الاعتيادية للمعاملات، فترات الثقة وقيم الاحتمال p لتحديد عما إذا كانت قيمة المعاملات تساوي صفراً. ويمكن اعتبار المعاملات مقاييس

لتأثير المتغير التفسيري على المتغير التابع، وفي هذه الحالة، ينبغي أن نكون في غاية الحذر فيما يختص معامل الوقت. فعلى سبيل المثال يمكننا تفسير نتائج مثل "يقيس β_2 تأثير المتغير التفسيري قبل فترتين على المتغير التابع مع بقاء العوامل الأخرى ثابتة". وباستثناء هذه الاختلافات الطفيفة، تعد الأساليب الإحصائية والتفسيرية مشابهة لتلك التي تم وصفها سابقاً. ورغم ذلك، فمن الضروري بحث هذه النوعية من النماذج بصوره منفصلة لأن من شأن ذلك مساعدتنا على فهم مصطلحات السلاسل الزمنية وطرح بعض الأفكار التي سنعتمد عليها في الفصول القادمة بشكل كبير.

وقبل تناول مثال توضيحي حول كيفية التعامل مع نماذج فترات التباطؤ الموزعة، دعنا نخرج قليلاً من الموضوع ونتناول نقطتين بإيجاز. تصف النقطة الأولى ماهية المتغيرات المتباطئة وكيفية حسابها في برنامج إكسل، في حين توضح الثانية الرموز المستخدمة والاختصارات في هذا الفصل والفصول التالية.

لمحة عن المتغيرات المتباطئة:

يعد مفهوم المتغيرات المتباطئة ضرورياً للتعامل مع بيانات السلاسل الزمنية، وسنتناول بالتفصيل معنى هذا المفهوم وكيفية إعداد المتغيرات المتباطئة والتعامل معها بواسطة برامج الحاسب الآلي.

افترض أن لدينا بيانات سلسلة زمنية لفترة زمنية t=1,...,T للمتغير X . وكما هو معتاد، سنرمز للمشاهدات الفردية ب(X_t) بالنسبة لـ= t=1,...,T .

انظر إمكانية استحداث متغير جديد W تتمثل مشاهداته في $W_t = X_t$ بالنسبة لـ: t=2,...,T ومتغيراً جديداً آخر Z تتمثل مشاهداته في $Z_t = X_{t-1}$ بالنسبة لـ: t=2,...,T. لماذا نكتب t=2,...,T بدلا عن t=1,...,T؟ إذا كتبنا: t=1,...,T فإن

المشاهدة الأولى في المتغير Z (أي Z_1) ستكون مساوية لـ: X_0. ولكننا لم نعرف بعد ما هو X_0 نظراً لأن X تتم مشاهدته فقط من: t=1,...,T. بعبارة أخرى، Z,W لهما T-1 مشاهدة فقط. لاحظ أيضاً أننا كتبنا $Z_t = X_{t-2}$ ومن ثم فإن المتغير الجديد Z سيكون له مشاهدات من t=3,...,T و T-2 مشاهدة فقط.

لكل من المتغيرين الجديدين Z,W عدد T-1 مشاهدة. فإذا تخيلنا أن Z,W عبارة عن عمودين يحتوي كل منهما على T-1 مشاهدة (كما في جداول برنامج إكسل)، فسوف نلاحظ أن العنصر الأول في W سيكون X_2 والعنصر الأول في Z سيكون X_1، وأن العنصر الثاني في W و Z سيكون X_2, X_3، ... الخ، ويمكن التعبير عن ذلك بالقول أن W يحتوي على X وأن Z يحتوي على X قبل فترة واحدة أو متباطئة فترة واحدة. عموما، يمكننا استحداث متغيرات "X متباطئ فترة واحدة"- أو "متباطئةX" اختصارا-"X متباطئ فترتين" أو بصفة عامة "X متباطئ عدد j من الفترات".

بإمكانك التأمل في "X كمتباطئة فترة واحدة"، "X كمتباطئة فترتين"،... الخ، كمتغيرات تفسيرية مختلفة تماماً كما الحال في "سعر المسكن"، "مساحة الأرض" أو "عدد غرف النوم" كمتغيرات تفسيرية مختلفة.

ورغم ذلك، لاحظ أنك إذا أردت إدراج متغيرات تفسيرية عديدة في نموذج الانحدار المتعدد، يجب أن تكون كافة المتغيرات متماثلة من حيث عدد المشاهدات. دعنا ننظر في معنى هذه العبارة في السياق التالي: افترض أن نموذج الانحدار يحتوي على: X يساوي سعر الفائدة متباطئاً j فترة كمتغير تفسيري. فإذا بدأت بـ: t=1,...,T من المشاهدات في سعر الفائدة، فإن X المتباطئة j فترة سوف يحتوي على j-T مشاهدة فقط. وبما أن هذا المتغير يحتوي على j-T مشاهدة فقط، يجب عليك التأكد من أن كافة المتغيرات الأخرى

في النموذج تحتوي تماماً على j -Tمشاهدة. بمعنى أن كل متغير في انحدار السلاسل الزمنية يجب أن يتضمن عدداً من المشاهدات يساوي T ناقصاً العدد الأقصى من المتباطئات الذي يتضمنه أي متغير.

تقوم بعض برامج الحاسب الآلي المتخصصة في الاقتصاد القياسي (مثل PcGive,SHAZAM,MicroFit) باستحداث المتغيرات المتباطئة تلقائياً بأمر بسيط ولكن برنامج إكسل لا يستطيع القيام بذلك. وهذا سبب رئيس لحاجتك إلى تعلم كيفية استخدام برامج الحاسب الآلي الأخرى المتخصصة في الاقتصاد القياسي مع بيانات السلاسل الزمنية بدلاً من استخدام برنامج إكسل. عند تعاملك مع برنامج إكسل يتعين عليك إنشاء المتغيرات المتباطئة بنفسك قبل البدء في إجراء الانحدار المتضمن لها. ومن شأن تقديم شرح مختصر لكيفية عمل ذلك أن يكون مفيداً عند تعاملك مع برنامج إكسل فضلاً عن توفير أسلوب عملي لتوضيح المادة المذكورة أعلاه.

افترض مثلاً أن لدينا 10 مشاهدات حول المتغيرين X,Y(t=1,...,10) وترغب في تقدير نموذج انحدار يتضمن X, X المتباطئ فترة واحدة، X المتباطئ فترتين، X المتباطئ ثلاث فترات، أي أننا نرغب في تقدير نموذج الانحدار:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 X_{t-2} + \beta_4 X_{t-3} + e_t.$$

يوضح الجدول رقم (8-1) كيفية ظهور البيانات في شكل جدول على برنامج إكسل.

لاحظ أن الجدول يحدد كل مشاهدة وفقاً للصف والعمود المعنيين، كما في الجدول رقم (8-1) يحتوي كل عمود على متغير (مثلاً، يحتوي العمود C على المتغير X المتباطئ فترة واحدة) ويحتوي كل صف على مشاهدات. لاحظ أن كل متغير يحتوي على 7 مشاهدات (T ناقصاً العدد الأقصى

للمتباطئات مثال: 10-3=7). فإذا نظرنا عبر أي صف (الصف 4 مثلاً) سنشاهد ما يلي: (أ) يتضمن Y و X بيانات عند نقطة زمنية محددة (مثلا، X_7, Y_7 أو $t = 7$)، (ب) سيحتوي X المتباطئ على مشاهدة من فترة زمنية واحدة في السابق (X_6 مثلاً)، (ج) سيحتوي X المتباطئ فترتين على مشاهدة من فترتين في السابق (X_5 مثلاً)، (د) سيحتوي X المتباطئ 3 فترات مشاهدة من 3 فترات سابقة (X_4 مثلاً).

بإمكانك إنشاء هذا الجدول على برنامج إكسل. استخدم أولاً أوامر قص/لصق في ملف إكسل المتضمن للبيانات الأصلية لكل من Y و X (أي الذي يحتوي على المشاهدات العشر الأصلية على المتغيرين) لإنشاء جدول إلكتروني يشابه الجدول رقم (8-1). بعد ذلك، نفذ عملية الانحدار باستخدام قائمة انحدار برنامج إكسل بالطريقة المعيارية وتحديد $A_1 : A_7$ في مربع "Input Y-range" و $B_1 : E_7$ في مربع "Input X-range".

يبدو هذا الجزء الخاص بالمتغيرات المتباطئة ذا علاقة محدودة بما يختص بفهم النتائج وتفسيرها. ولكن من المهم تذكر هذه المادة عند التعامل عملياً مع بيانات السلاسل الزمنية بواسطة الحاسب الآلي.

جدول رقم (8-1): إنشاء المتغيرات المتباطئة

	العمود A (Column A) Y	العمود B (Column B) X	العمود C X المتباطئ فترة واحدة (Column C X Lagged One period)	العمود D X المتباطئ فترتان (Column D X Lagged two periods)	العمود E X المتباطئ ثلاث فترات (Column E X Lagged Three periods)
الصف1(Row1)	Y_4	X_4	X_3	X_2	X_1
الصف2(Row2)	Y_5	X_5	X_4	X_3	X_2
الصف3(Row3)	Y_6	X_6	X_5	X_4	X_3
الصف4(Row4)	Y7	X_7	X_6	X_5	X_4
الصف5(Row5)	Y_8	X_8	X_7	X_6	X_5
الصف6(Row6)	Y_9	X_9	X_8	X_7	X_6
الصف7(Row7)	Y_{10}	X_{10}	X_9	X_8	X_7

لمحة عن الرموز:

من الأهمية بمكان أيضاً التأكد من أن الرموز واضحة تماماً. خذ المتغير X (الكثافة السكانية، مثلاً). بعد جمع البيانات عن X سنحصل على المشاهدة X_i بالنسبة لـ: i=1,....,N في حالة البيانات المقطعية المختلفة و X_t بالنسبة لـ: t=1,...,T في حالة بيانات السلاسل الزمنية (انظر الفصل الثاني).

بعبارة أخرى، X عبارة عن رمز عام للمتغير و X_i أو X_t يشير إلى مشاهدة محددة للمتغير (مثلاً: X_i =الكثافة السكانية في الدولة رقم i أو X_t =الكثافة السكانية في الفترة الزمنية t). في نقاشنا لموضوع الانحدار في الفصول (4-7)، كنا غالباً نستخدم معادلات من نوع:

$$Y = \alpha + \beta X + e.$$

ويمكن التعبير عن هذه المعادلة بالعبارة التالية: "يعتمد المتغير التابع Y على المتغير التفسيري X في علاقة خطية". وعند حصولنا على البيانات الفعلية، يمكن كتابة المعادلة كما يلي:

$$Y_i = \alpha + \beta X_i + e_i.$$

ويمكن التعبير عنها بالعبارة التالية: "تعتمد قيمة المشاهدة i من المتغير Y على قيمة المشاهدة i من المتغير X". فعلى سبيل المثال، يعتمد انحسار مساحة الغابات في الدولة i على الكثافة السكانية في الدولة i". والمعادلتان صحيحتان تماماً، ولكن لما كان الرمز السفلي i في المعادلة الأخيرة أمراً بديهياً (أي من البديهي أن يعتمد انحسار مساحة الغابات في جمايكا على الكثافة السكانية في جمايكا وليس أوغندا)، فغالباً ما يتم إسقاط الرمز i من المعادلة الثانية لغرض التبسيط.

لقد استخدمنا الرموز على نحو أكثر في الفصل السادس عند مناقشة الانحدار المتعدد، إذ كانت المتغيرات $X_1, X_2, ..., X_k$ عبارة عن κ من المتغيرات التفسيرية المختلفة. هنا يشير الرمز الموجود على X إلى أي المتغيرات التفسيرية (وليس المشاهدة) هو المعني. وقد حاولنا في مرات قليلة جداً أن نكون أكثر وضوحاً إذ كتبنا X_{2i} ، مثلا، للإشارة إلى المشاهدة i من المتغير التفسيري الثاني.

ورغم ذلك، ولما كان من البديهي عادة في حالة الانحدار المتعدد أن يكون Y_i (انحسار مساحة الغابات في الدولة i) معتمداً على X_{1i} (الكثافة السكانية في الدولة i) وعلى X_{2i} (التغير في المراعي في الدولة i)، فإنه غالباً ما يتم إسقاط الرمز i من المعادلة.

باختصار، يمكن القول بأن الرموز المستخدمة في هذا الكتاب للتمييز بين متغير ومشاهدة محددة لهذا المتغير تتسم بقدر من المرونة. وهذا أمر متبع (وشائع في الكتب المنهجية) مادام المعنى واضحاً من السياق وإلا لأصبحت المعادلات تتضمن عدداً هائلاً من الرموز. سنوضح في فصول السلاسل الزمنية في هذا الكتاب كيف أننا سنستخدم X_{t-j} للإشارة إلى مشاهدة معينة (مثلا: إذا كان t=1968 و j=3، فإن X_{t-j} هو قيمة المتغير X في سنة 1965) وأن المتغير X متباطئ بمقدار j فترة زمنية، وهذا واضح من السياق. حقيقة، فيما يخص أى معادلة وردت في هذا الكتاب لايهم على أى نحو يتم تفسير تلك المعادلة.

مثال: تأثير التدريب في مجال السلامة على عدد الحوادث:

من الممكن أن تصل الخسائر التي تتحملها الشركات الكبرى نتيجة للحوادث الصناعية أرقاماً كبيرة جداً. لذلك، تعمل العديد من الشركات على توفير التدريب الكافي في مجال السلامة للعاملين بها بهدف الحد من تلك الحوادث. وعادة تكون هذه الشركات حريصة على معرفة مدى فعالية هذه البرامج التدريبية.

يحتوى ملف إكسل (SAFETY.XLS) على بيانات عن شركة معينة تم جمعها على أساس شهري خلال خمس سنوات (60 شهراً) حول المتغيرات التالية:

Y= الخسائر الناجمة عن الحوادث (جنيه إسترليني في الشهر).

X= عدد ساعات التدريب في مجال السلامة لكل عامل شهرياً.

ولما كانت هذه البيانات عبارة عن سلسلة زمنية وأنه من المتوقع أن يؤثر التدريب في مجال السلامة في الشهر السابق على معدلات الحوادث الحالية، فمن الضروري إدراج متباطئات X في الانحدار. يتضمن الجدول رقم (8-2) تقديرات مربعات صغرى اعتيادية للمعاملات في نموذج متباطئات موزعة يتم فيها السماح باعتماد الخسائر على التدريب الحالي في مجال السلامة وعلى التدريب في مجال السلامة خلال الشهور الأربعة السابقة. وبذلك، نحصل على المعادلة التالية:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + e_t.$$

ما هو استنتاج الشركة بشأن فعالية برامج التدريب؟ ستكون زيادة عدد ساعات التدريب في مجال السلامة لكل عامل بمقدار ساعة واحدة في الشهر مصحوبة بالآتي:

1- انخفاض فوري في الخسائر الناجمة عن الحوادث بمقدار 145.00 جنيه إسترليني مع بقاء العوامل الأخرى ثابتة.

2- انخفاض الخسائر بمقدار 462.14 جنيهاً إسترلينياً بعد شهر مع بقاء العوامل الأخرى ثابتة.

3- انخفاض الخسائر بمقدار 424.47 جنيهاً إسترلينياً بعد شهرين مع بقاء العوامل الأخرى ثابتة.

جدول رقم (8-2): نتائج التقدير باستخدام نموذج فترات التباطؤ الموزعة لتأثير التدريب في مجال السلامة على الحوادث

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
القاطع (Intercept)	92001.51	2001.71	45.96	5.86E-42	87978.91	96024.11
X_t	-145.00	47.62	-3.04	0.0037	-240.70	-49.30
X_{t-1}	-462.14	47.66	-9.70	5.52E-13	-557.91	-366.38
X_{t-2}	-424.47	46.21	-9.19	3.12E-12	-517.33	-331.62
X_{t-3}	-199.55	47.76	-4.18	0.00012	-295.52	-103.58
X_{t-4}	-36.90	47.45	-0.78	0.44	-132.25	58.45

4- انخفاض الخسائر بمقدار 199.55جنيهاً إسترلينياً بعد 3 أشهر مع بقاء العوامل الأخرى ثابتة.

5- انخفاض الخسائر بمقدار 36.9 جنيهاً إسترلينياً بعد 4 أشهر مع بقاء العوامل الأخرى ثابتة.

بالإمكان تفسير فترات الثقة حسب الطريقة المعتادة. فعلى سبيل المثال، فنحن واثقون بدرجة 95% من الانخفاض الفوري في الخسائر الناجمة عن الحوادث بمقدار 49.30 جنيهاً إسترلينياً على الأقل و 240.70 جنيهاً إسترلينياً على الأكثر مع بقاء العوامل الأخرى دون تغيير.

لمعرفة ماذا تعنى عبارة بقاء العوامل الأخرى دون تغيير (ثابتة) في هذا السياق، لاحظ أنه يمكننا التعبير عن العبارة الثانية، مثلاً، على النحو التالي:

"تميل زيادة التدريب في مجال السلامة خلال شهر معين إلى تخفيض خسائر الحوادث في الشهر التالي بمقدار 462.14 جنيهاً إسترلينياً مع افتراض عدم حدوث أي تغيير في سياسة الشركة المتعلقة بالتدريب في مجال السلامة".

عند فحص النتائج الإحصائية للجدول رقم (8-2)، نلاحظ أن كافة المعاملات ذات معنوية إحصائياً باستثناء β_4 حيث تبلغ قيمة P لهذا المعامل 0.44 (ليس أقل من 0.05). كما نلاحظ أن فترة الثقة للمعامل β_4 تتضمن الصفر. ونتيجة لذلك لا يمكننا رفض فرضية أن $\beta_4 = 0$. لذلك، لا يمكننا رفض فرضية أن التدريب في مجال السلامة قبل أربعة أشهر سابقة لا يؤثر في الخسائر الحالية الناجمة عن الحوادث. فذلك يعني أن العاملين نسوا ماتدربوا عليه بعد مضي أربعة أشهر. وهذه معلومة مفيدة للشركة لأنها قد تقترح إجراء برامج التدريب في مجال السلامة مرة كل أربعة أشهر.

عموماً، يظهر تأثير التدريب في مجال السلامة على الحوادث نمطاً أشبه بشكل هرمي عبر الزمن: يعد التأثير الفوري للتدريب في مجال السلامة على الحوادث صغيراً جداً (145 جنيهاً إسترلينياً). ولكن التأثير يتزايد بعد ذلك لأكثر من 400 جنيه إسترليني لكل من الشهرين التاليين، ثم ينخفض إلى نحو 200 جنيه بعد ثلاثة أشهر وينخفض إلى صفر جنيه تقريباً بعد أربعة أشهر. لاحظ أن زيادة التدريب في مجال السلامة في فترة معينة تميل إلى تخفيض الخسائر فوراً وخلال الفترات القليلة التالية أيضاً. فإذا جمعنا الفوائد الناجمة عن زيادة التدريب في مجال السلامة بمقدار ساعة واحدة في كل فترة (أي 145ج+462.14ج+424.47ج+199.55ج+36.90ج=1268.06ج) [2]، فإننا نحصل على مقياس للفوائد الإجمالية لبرنامج التدريب في مجال السلامة. بعبارة أخرى، يمكن القول بأن "إضافة ساعة واحدة من التدريب في مجال السلامة تميل إلى تحقيق منافع إجمالية مقدارها 1268.06جنيهاً في الشهر الذي تم فيه التدريب وفي الشهور الأربعة التالية".

بحساب هذه الفوائد الإجمالية والنظر إلى نمط المعاملات عبر الزمن، تحصل الشركة على معلومات قيمة يمكن استخدامها في إعادة تصميم برامجها التدريبية المتعلقة بالسلامة. ورغم ذلك، فإن هذه النتائج تفترض أن نموذج فترات التباطؤ الموزعة لايتجاهل أي متغير تفسيري آخر. فعلى سبيل المثال، إننا نفترض بوضوح أن X_{t-5} (التدريب في مجال السلامة قبل خمسة أشهر) ليس له تأثيرعلى الخسائر الحالية الناجمة عن الحوادث. فإذا كان هذا الافتراض غير صحيح، فإن تقديراتنا فيما يختص بفائدة التدريب في مجال السلامة سوف تكون غير صحيحة. وهذه القضية وثيقة الصلة بمشكلة تحيز المتغيرات المحذوفة التي تمت مناقشتها في الفصل السادس. كما أنها تؤكد أهمية الاختيار الصحيح لطول المتباطئة (q في نموذج فترات التباطؤ الموزعة). وهذا هو الموضوع الذي سنتناوله لاحقاً.

تمرين رقم (8-1):

استخدم ملف البيانات SAFETY.XLS، الواردة في المثال السابق في حل هذا التمرين. تتضمن منظومة البيانات ما يلي: T=60 مشاهدة في Y= خسائر الحوادث و X= عدد ساعات التدريب في مجال السلامة.

أ- قم بإنشاء المتغيرات التفسيرية التي ستستخدمها في نموذج فترات التباطؤ الموزعة مع استخدام طول متباطئة مقداره 4. (إذا واجهتك مشكلة انظر في الملف SAFETY.XLS الذي يتضمن الإجابة عن هذا السؤال).

كم عدد المشاهدات في المتغيرات التفسيرية؟

ب- باستخدام إجابة الفقرة (أ)، قم بإعادة تكوين الجدول في المثال أعلاه.

جـ- قم بإنشاء المتغيرات التفسيرية التي ستستخدمها في نموذج فترات تباطؤ موزعة يبلغ طول المتباطئة فيه 2. كم عدد المشاهدات في المتغيرات التفسيرية؟

د- باستخدام إجابة الفقرة (جـ)، قدر نموذج فترات التباطؤ موزعة حيث q=2.

هـ- قارن إجابة الفقرة (د) مع إجابة الفقرة (ب) ووضح أسباب الاختلاف مع الاهتمام بصفة خاصة بمسألة تحيز المتغيرات المحذوفة (انظر الفصل السادس إذا كنت قد نسيت ذلك).

اختيار درجة التباطؤ:

عند التعامل مع نماذج فترات التباطؤ الموزعة، نادراً ما نعرف على وجه التحديد كم عدد المتباطئات التي يجب إدراجها في النموذج. لماذا افترضنا، في النموذج السابق، أن الخسائر تعتمد على التدريب في مجال السلامة حتى أربعة أشهر سابقة؟ لماذا لم نفترض ثلاثة أو ستة أو حتى ثمانية أشهر؟ السبب هو أننا وبعكس معظم نماذج الانحدار التي تم بحثها في الفصول (4-7)، لا نعرف أي المتغيرات التفسيرية في نموذج فترات تباطؤ موزعة يدخل في الانحدار قبل أن نجلس عملياً أمام الحاسب الآلي ونبدأ في معالجة البيانات. لذلك، من المناسب القول بأن مشكلة اختيار طول المتباطئة تصبح مشكلة بيانات حيث يتم استخدام المتوسطات الإحصائية لتحديد عدد المتباطئات المطلوب إدراجها.

هناك طرق مختلفة لاختيار طول المتباطئة في مجالات الاقتصاد القياسي. وسوف نوضح هنا طريقة شائعة ولا تتطلب أي أساليب إحصائية جديدة بخلاف

تلك التي تعرفنا عليها في الفصل الخامس. تستخدم هذه الطريقة اختبار t لمعرفة ما إذا كان $\beta_q = 0$ بهدف تحديد طول المتباطئة. إن الطريقة الأكثر استخداماً هنا هي:

(أ) البدء بطول متباطئة كبير [3] ($B_{q_{\max}}$) وإجراء اختبار لمعرفة ما إذا كان معامل المتباطئة العظمى يساوي صفراً (أي $B_{q_{\max}} = 0$).

(ب) إذا كان يساوي صفراً، اسقط هذه المتباطئة العظمى وأعد تقدير النموذج باستخدام متباطئة عليا تساوي $q_{\max-1}$.

(ج) فإذا اتضح أن $B_{q_{\max-1}} = 0$ في الانحدار الجديد، قم بتخفيض رتبة المتباطئة بمقدار الواحد ثم أعد تقدير النموذج.

(د) استمر في تخفيض رتبة المتباطئة بمقدار الواحد وإعادة تقدير النموذج حتى يتم رفض "فرضية أن معامل المتباطئة العظمى يساوي صفراً".

ولجعل أسلوب اختيار المتباطئة هذا أسلوباً منتظماً نتبع الخطوات التالية:

الخطوة الأولى: اختر أكبر طول ممكن للمتباطئة ($q_{\max}$)حسب تقديرك.

الخطوة الثانية: قدر نموذج فترات التباطؤ الموزعة:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_{q_{\max}} X_{t-q_{\max}} + e_t.$$

فإذا كانت قيمة p عند اختبار فرضية أن $\beta_{q_{\max}} = 0$ أقل من مستوى المعنوية الذي اخترته (0.05 مثلاً) فلا تستمر في العملية واستخدام $q_{\max}$ كطول للمتباطئة، وإلا انتقل للخطوة التالية.

الخطوة الثالثة: قدر نموذج فترات التباطؤ الموزعة:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_{q_{\max}-1} X_{t-q_{\max}+1} + e_t.$$

فإذا كانت قيمة p عند اختيار فرضية أن $\beta_{q_{\max}-1} = 0$ أقل من مستوى المعنوية الذي اخترته (0.05 مثلاً)، فلا تستمر في العملية واستخدم $q_{\max-1}$ كطول للمتباطئة، وإلا فاستمر للخطوة التالية .

الخطوة الرابعة: قدر نموذج فترات التباطؤ الموزعة:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_{q_{\max}-2} X_{t-q_{\max}+2} + e_t.$$

فإذا كانت قيمة p عند اختيار فرضية $\beta_{q_{\max-2}} = 0$ أقل من مستوى المعنوية الذي اخترته (0.05 مثلاً)، فلا تستمر في العملية واختر $q_{\max} - 2$ كطول للمتباطئة، وإلا فاذهب للخطوة التالية،... الخ.

من الملائم عملياً، عند تعاملك مع الجداول الإحصائية، ملاحظة أن عدد المشاهدات المستخدمة في نموذج فترات التباطؤ الموزعة يساوي العدد الأصلي للمشاهدات T ناقصاً طول المتباطئة الأقصى. وهذا يعني أننا في الخطوة (2) نتعامل مع ($T - q_{\max}$) من المشاهدات، وفي الخطوة (3) مع ($T - q_{\max} + 1$) من المشاهدات، وفي الخطوة (4) مع ($T - q_{\max} + 2$) من المشاهدات، وهكذا. فكل خطوة في حاجة إلى بعض القص واللصق في الجدول الإحصائي لبرنامج إكسل لاستحداث المتغيرات مع العدد المناسب من المشاهدات[4].

مثال: تأثير التدريب في مجال السلامة على الحوادث: (تتمة ما ورد في الصفحات السابقة)

افترض أن ($q_{max} = 4$) في انحدار فترات التباطؤ الموزعة للخسائر الناجمة عن الحوادث في التدريب على وسائل السلامة. بعبارة أخرى، إننا نعتقد أن فترة الأربعة أشهر هي الفترة القصوى التي نتوقع فيها تأثير للتدريب في مجال السلامة على الخسائر الناجمة عن الحوادث. تشير الطريقة الموضحة أعلاه إلى أنه يجب علينا البدء بتقدير نموذج فترات التباطؤ الموزعة بطول متباطئة مقداره 4. يوضح الجدول السابق النتائج. وبما أن قيمة p المقابلة للمتغير التفسيري X_{t-4} أكبر من 0.05 ، فإننا لا نستطيع رفض فرضية أن $\beta_4 = 0$ عند درجة ثقة 5%. لذلك، نسقط هذا المتغير من النموذج ونعيد تقدير طول المتباطئة ليصبح 3 وبالتالي الحصول على النتائج الموضحة في الجدول رقم (8-3).

جدول رقم (8-3): تحديد طول متباطئة عند 3 فترات

	معامل الانحدار (Coefficient)	الخطأ المعياري (Standard Error)	إحصائية t (t-stat)	قيمة الاحتمال (p-value)	الحد الأدنى لفترة الثقة 95% (Lower 95%)	الحد الأعلى لفترة الثقة 95% (Upper 95%)
القاطع (Inercept)	90402.22	1643.18	55.02	9.19E-48	87104.94	93699.51
X_t	-125,90	46.24	-2.72	0.0088	-218.69	-33.11
X_{t-1}	-443.49	45.88	-9.67	3.32E-13	-535.56	-351.42
X_{t-2}	-417.61	45.73	-9.13	2.18E-12	-509.38	-325.84
X_{t-3}	-179.90	46.25	-3.89	0.0003	-272.72	-87.09

إن قيمة P المطلوبة لاختبار فرضية $\beta_3 = 0$، هي 0.0003، وهي تقل كثيراً عن 0.05، ومن ثم نستنتج أن المتغير X_{t-3} متغير معنوي إحصائياً في نموذج فترات التباطؤ الموزعة؛ لذا فإن q=3 هو طول المتباطئة المختار لهذا النموذج. وبما أن هذه النتائج مماثلة لتلك التي تم بحثها أعلاه، فلا داعي لإعادة تفسيرها.

تمرين رقم (8-2):

استخدام ملف البيانات، SAFETY.XLS، الذي يحتوي على 60 مشاهدة (T=60) للمتغيرات Y = خسائر الحوادث و X = عدد ساعات التدريب في مجال السلامة. افترض أنك تعتقد أن فترة ستة أشهر هي الفترة القصوى التي قد يؤثر فيها التدريب في مجال السلامة على خسائر الحوادث؛ لذا فإنك تختار $q_{\max} = 6$. باستخدام الطريقة الموضحة أعلاه، اختر طول المتباطئة المناسب لنموذج المتباطئات الموزعة.

تمرين رقم (8-3):

يهتم المختصون في مجال اقتصاديات التنمية في الغالب بدراسة أثر الإنفاق على التعليم على النمو الاقتصادي. ومع ذلك، يساورهم الشك أحياناً في أن التأثيرات الإيجابية لتحسين المستوى التعليمي قد تظهر بعد 5 أو 10 سنوات في شكل معدلات نمو أعلى . وفي ضوء هذه الاعتبارات، استخدم ملف البيانات التالي في تقدير نموذج وإعداد تقرير موجز حول تأثير زيادة الإنفاق على التعليم الأساسي على النمو الاقتصادي.

يتضمن ملف البيانات (EDUC.XLS) البيانات السنوية لأحد الدول من عام 1910 حتى عام 1995 حول المتغيرات التالية:

Y= نمو إجمالي الناتج المحلي (GDP) كنسبه مئوية في السنة.

X= الإنفاق على التعليم (بالدولار لكل طفل دون سن 16 سنة).

ملخص الفصل:

1) تتضمن معادلات الانحدار مع متغيرات السلاسل الزمنية مشكلتين لم نتعامل معهما في السابق وهما: أولاً، بإمكان أي متغير التأثير في متغير آخر بفترات تباطؤ. ثانياً، إذا كانت المتغيرات غير مستقرة، فمن المحتمل حدوث مشكلة الانحدار الزائف. سيتم بحث هذه القضية في الفصل العاشر.

2) يعتمد المتغير التابع، في نماذج فترات التباطؤ الموزعة، على المتغير التفسيري وفترات التباطؤ للمتغير التفسيري.

3) إذا كانت المتغيرات في نموذج فترات التباطؤ الموزعة مستقرة، تكون تقديرات المربعات الصغرى الاعتيادية موثوقة ويصبح بالإمكان استخدام الأساليب الإحصائية للانحدار المتعدد (مثلاً، اختبار قيم p أو فترات الثقة) بصورة مباشرة.

4) يمكن اختيار طول المتباطئة في نموذج فترات تباطؤ موزعة بالاستخدام المتتابع لاختبارات t ابتداء بطول متباطئة على قدر مناسب من الطول.

ملحق رقم (8-1): نماذج فترات تباطؤ موزعة أخرى:

يعد نموذج فترات التباطؤ الموزعة، الذي يتناوله هذا الفصل، نموذجا عاماً إلى حد بعيد. فليست هناك تقييدات على القيم التي تأخذها المعاملات $\beta_0, \beta_1, ..., \beta_q$. ومع ذلك، هناك العديد من النماذج الأخرى في مجالات الاقتصاد القياسي التي تعتبر نماذج فترات تباطؤ موزعة ولكنها تفرض قيوداً على المعاملات. ولأن التعامل مع هذه النماذج يعد صعباً بعض الشيء (على الأقل في برنامج إكسل)، فسوف نتناولها بإيجاز في شكل ملحق.

هناك العديد من النماذج التي تفرض قيوداً على نماذج فترات التباطؤ الموزعة (مثلاً، نموذج المتباطئة الحسابية، نموذج المتباطئة الهندسية، نموذج Koych،... الخ)، ولكننا لن نناقشها بالتفصيل. ويمكننا أخذ نموذج فترات التباطؤ الموزعة ذات الحدود المتعددة (Polynomial Distributed Lag Model) (متباطئة Almon) كمثال لهذه الأنواع من النماذج واستخدامها لتوضيح القضايا الأساسية ذات الصلة.

إن نموذج فترات التباطؤ الموزعة ذات الحدود المتعددة هو ذاته نموذج المتباطئة الموزعة مع إضافة التقييد التالي:

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_i i^2$$

أي أن معاملات المتباطئة الموزعة مقيدة كدالة من الدرجة الثانية لطول المتباطئة (تعد الصيغة التربيعية الخيار الأفضل للحدود المتعددة في هذا النموذج الخاص بالمتباطئات الموزعة. ولكن هناك صيغ أخرى ممكنة مثل التكعيبية). تعتمد هذه الدالة التربيعية (من الدرجة الثانية) على ثلاثة معاملات مجهولة ($\gamma_0, \gamma_1, \gamma_2$) يجب تقديرها. لاحظ أننا بمجرد حصولنا على تقديرات Y_0, Y_1, Y_2 يمكننا استخدامها مع المعادلة أعلاه للحصول على تقديرات β_i لكل من $i = 0,1,....., q$. وبمجرد حصولنا على تقديرات للمعاملات الأخيرة يصبح من الممكن تفسيرها بالطريقة ذاتها المستخدمة في السابق.

كيف يمكننا تقدير Y_0, Y_1, Y_2؟ بإمكاننا إجراء انحدار مربعات صغرى اعتيادية ولكن باستخدام متغيرات تفسيرية غير عادية. سنوضح ذلك باستخدام q=3. سيكون نموذج فترات التباطؤ الموزعة كما يلي:

$$. Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + e_t$$

فإذا استبدلنا $\beta_i s$ باستخدام المعادلة التربيعية أعلاه، يمكننا كتابة نموذج فترات التباطؤ الموزعة كثيرة الحدود على النحو التالي:

$$Y_t = \alpha + \gamma_0 V_t + \gamma_1 W_t + \gamma_2 Z_t + e_t$$

حيث إن:

$$V_t = X_t + X_{t-1} + X_{t-2} + X_{t-3}$$

$$W_t = X_{t-1} + 2X_{t-2} + 3X_{t-3}$$

$$Z_t = X_{t-1} + 4X_{t-2} + 9X_{t-3}$$

بعبارة أخرى، يمكننا الحصول على تقديرات المربعات الصغرى الاعتيادية لكل من $\gamma_0, \gamma_1, \gamma_2$ بإجراء معادلة الانحدار Y للمتغيرات Z,W,V. يجب إنشاء المتغيرات التفسيرية V و Z,W لمعادلة الانحدار هذه.

لقد عرفنا الآن ماهو نموذج المتباطئات الموزعة كثيرة الحدود وكيفية تقديره. ولكن ربما يكون من غير الواضح سبب استخدام مثل هذا التقدير. هناك سببان وثيقان لدراسة إمكانية فرض هذه القيود متعددة الحدود على نموذج فترات التباطؤ الموزعة:

1- هناك عدد أقل من المعاملات المطلوب تقديرها لنموذج فترات التباطؤ الموزعة. ففي المعادلة التربيعية أعلاه، يبلغ عدد المعاملات عادة 3 معاملات ($\gamma_0, \gamma_1, \gamma_2$). وسيكون لنموذج فترات التباطؤ الموزعة q+1 معامل. وفي الواقع فإن q قد يكون كبيراً جداً (مثلا، قد تختار q=12 إذا كانت لديك بيانات شهرية وتريد فترة تباطؤ تمتد لسنة). وقد يكون من

الصعب الحصول على تقديرات موثوقة لعدد كبير من المعاملات ما لم يكن عدد المشاهدات كبيراً جداً.

2- قد تعانى نماذج المتباطئات الموزعة أحياناً من الارتباط الخطي المتعدد (انظر الفصل السادس). فعلى سبيل المثال، فإن X_t و X_{t-1} قد يكونان مرتبطين بصورة وثيقة. افترض، مثلاً، أن X= سعر الفائدة. وبما أن أسعار الفائدة تتغير عادة ببطء شديد عبر الزمن، فغالباً ما يكون X_t و X_{t-1} متساويين (أو مطابقين) لبعضهما، الأمر الذي يجعلهما مرتبطين بصورة وثيقة ومن ثم بروز مشكلة الارتباط الخطي المتعدد. إن نماذج فترات التباطؤ الموزعة متعددة الحدود لا تعاني عادة مثل هذه المشكلات (مثلاً، لا تكون Z,W,V في المثال السابق مرتبطة بصورة وثيقة).

السبب الأخير لا يكون مشكلة في العديد من الحالات بالنسبة للتحليل، ولكن إذا حدث أن أصبح مشكلة، فغالباً ما يرجع السبب إلى أن المتغيرات التفسيرية غير مستقرة. سنبحث موضوع المتغيرات التفسيرية غير المستقرة في الفصل العاشر ونوضح طريقة حل مشكلة الارتباط الخطي المتعدد في هذه الحالة دون الحاجة إلى فرض قيود على المعاملات.

إن المبرر الأول لاستخدام نماذج فترات التباطؤ الموزعة ذات الحدود المتعددة قد يكون ملزماً إذا كان عدد المشاهدات صغيراً جداً. ولكن مع كثرة البيانات المتاحة لخبراء الاقتصاد الكلي حالياً، يفقد هذا المبرر أهميته. علاوة على ذلك، فإن فرض قيود على المعاملات قد يؤدى إلى نتائج مضللة إذا كانت هذه القيود خاطئة. تذكر أن β_i يقيس تأثير المتغير التفسيري لعدد i فترة سابقة على القيمة الحالية للمتغير التابع. وقد يكون مثل هذا التأثير كبيراً في نموذج فترات التباطؤ الموزعة. يوضح نموذج فترات التباطؤ الموزعة كثيرة الحدود أن

S β_i يجب أن تكون ذات علاقة تربيعية مع بعضها. فإذا لم تكن العلاقة كذلك، فإن نتائج نموذج فترات التباطؤ الموزعة ذات الحدود المتعددة قد تكون مضللة إلى حد بعيد.

عموماً، يمكن القول أنه نادراً ما تكون هناك حالة ملحة لفرض قيود على نماذج فترات التباطؤ الموزعة؛ ولهذا السبب لن نناقشها في هذا الكتاب.

ملاحظات ختامية:

1- بإمكاننا ترميز معاملاتنا بالطريقة التي نريدها. إن الطريقة المختارة هنا تربط الرمز الموضوع على β بعدد الفترات السابقة التي تشير إليها المتغيرات التفسيرية. فعلى سبيل المثال، β_1 هو المعامل الخاص بالمتغير التفسيري X_{t-1} ، أي قيمة المتغير التفسيري مع فترة تباطؤ واحدة.

2- يمثل مبلغ 1268.06جنيهاً إسترلينياً تقديرات المبلغ الإجمالي للتوفير. وبالإمكان حساب فترة الثقة أيضاً ولكن سيتطلب ذلك معادلة أكثر تعقيداً وخارج نطاق هذا الكتاب.

3- تذكر أن كل متغير في نموذج فترات التباطؤ الموزعة سيكون له عدد من المشاهدات (وإن كان غير كبير جداً) يساوي T ناقصاً عدد المتباطئات الأقصى. فإذا اخترت رقماً كبيراً جداً لعدد المتباطئات الأقصى، فإن عدد المشاهدات سيكون قليلاً جداً.

4- كبديل لذلك، يستخدم بعض الباحثين بسهولة $T-q_{\max}$ مشاهدة لكافة أنواع الانحدار. ومن ميزات هذا البديل أن الباحث يستخدم، في كل خطوة، المشاهدات ذاتها. ورغم ذلك، فإن هذه الطريقة قد تعني استخدام بيانات أقل من المطلوب. تذكر(انظر الفصل الخامس) أن استخدام مشاهدات أكثر يزيد من درجة دقة تقديرات المربعات الصغرى الاعتيادية.

الفصل التاســع

تحليل السلاسل الزمنية الأحادية

(Univariate Time Series Analysis)

لقد ناقشنا في الفصل السابق نماذج فترات التباطؤ الموزعة، وهي تمثل نوعاً مبسطاً من نماذج الانحدار تم تصميمه للاستخدام مع بيانات السلاسل الزمنية. تذكر أن النموذج يفترض اعتماد المتغير التابع Y_t على متغير تفسيري X_t ومتباطئات المتغير التفسيري $x_{t-1},......,x_{t-q}$. حيث تمثل مثل هذه النماذج خطوة أولى مهمة في استيعاب المفاهيم الأساسية لتحليل السلاسل الزمنية.

بالإمكان استخدام نماذج فترات التباطؤ الموزعة، في العديد من الحالات، بدون مشكلات تذكر. ورغم ذلك، فإنها قد تكون مضللة في حالات معينة على سبيل المثال:

(1) عند اعتماد المتغير التابع Y_t على متباطئات المتغير التابع أيضاً، ربما في شكل x_t، $x_{t-1},......,x_{t-q}$.

(2) عندما تكون المتغيرات "غير مستقرة" (nonstationary).

وفقاً لذلك، سيتم في هذا الفصل استخدام أدوات للتعامل مع الحالتين السابقتين وتعريف معنى مصطلح "غير مستقرة". ولتبسيط التحليل، سنتجاهل X في هذا الفصل ونركز كلياً على Y. ومن حيث المصطلحات الإحصائية، سنركز على طرق السلاسل الزمنية الأحادية. وكما هو واضح من الاسم، فإن هذه الطرق تختص بمتغير واحد، أو بلغة الإحصاء سلسلة واحدة (مثال: Y= إجمالي الناتج المحلي الحقيقي). ومن الأهمية بمكان في هذا الصدد، فهم خصائص كل سلسلة على حدة قبل الإقدام على استخدام نماذج الانحدار التي تتضمن سلاسل زمنية متعددة.

مثال: الدخل الشخصي في الولايات المتحدة:

يمثل الشكل رقم (9-1) رسماً بيانياً لسلسلة واحدة: اللوغاريثم الطبيعي للدخل الشخصي (Personal Income) في الولايات المتحدة من الربع الأول لعام 1954 حتى الربع الأخير من عام 1994 [1]. بعبارة أخرى، Y_t هو الدخل الشخصي بالنسبة لـ: t=1954Q1,....,1994Q4. وهذه البيانات متاحة في ملف إكسل: INCOME.XLS. يتم قياس متغير الدخل الشخصي الأصلي بملايين الدولارات.

لاحظ أن الدخل الشخصي يبدو متزايداً مع مرور الزمن بمعدل ثابت تقريباً. وبالإمكان ملاحظة بعض التباين (مثل الانخفاض البسيط في الدخل الشخصي في فترات الكساد التي حدثت في منتصف السبعينيات وأوائل الثمانينيات) ولكن بصفة عامة يوضح الرسم البياني للسلسلة الزمنية خطاً مستقيماً بدرجة ميل موجبة. إن التحرك المستمر نحو الصعود يعرف بالاتجاه العام (Trend). وتُظهر متغيرات اقتصادية كلية عديدة (مثل إجمالي الناتج المحلي، مستوى الأسعار، الإنتاج الصناعي، الاستهلاك، الإنفاق الحكومي، ... الخ) اتجاهات عامة مماثلة.

من الملائم، عند هذه النقطة، التعريف بمفهوم الفروقات (Differencing). فإذا كان $Y_t (t=1,....,T)$ عبارة عن متغير سلسلة زمنية، فإن $\Delta Y_t = Y_t - Y_{t-1}$ يمثل الفرق الأول لـ: Y_t [2]. يقيس ΔY_t التغير في المتغير أو نموه عبر الزمن. فإذا أخذنا اللوغاريثمات الطبيعية للسلسلة الأصلية (Y_t)، فإن ΔY_t سوف يقيس التغير النسبي في المتغير

بين الفترة الزمنية t-1 والفترة الزمنية t. ويطلق على ΔY_t غالباً اسم ΔY أو دلتا Y أو "التغير في Y". علاوة على ذلك، كثيراً ما يشار إلى Y_{t-1} على أنها Y_t متباطئة فترة واحدة. يوضح الشكل رقم (9-2) التغير في الدخل الشخصي باستخدام البيانات المتاحة في ملف INCOME.XLS.

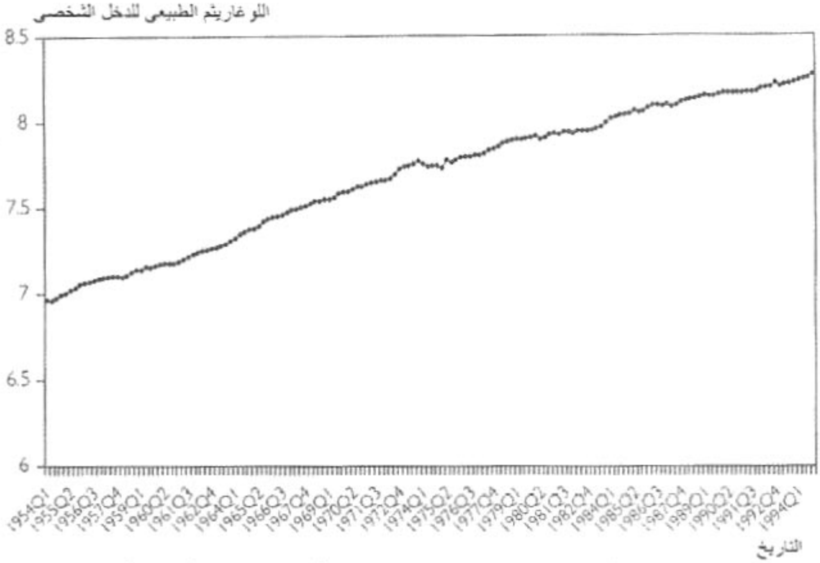
لاحظ أن الشكل رقم (9-2) يبدو مختلفاً تماماً عن الشكل رقم (9-1). لقد اختفى السلوك الاتجاهى الظاهر في الشكل رقم (9-1) تماماً (سنعود إلى هذه النقطة لاحقاً). يوضح الشكل ميل الدخل الشخصي إلى الارتفاع بنحو 1% في كل ربع سنة على الرغم من وجود قدر كبير من التباين في معدل النمو هذا عبر الزمن. ينخفض الدخل الشخصي في بعض الفترات الانكماشية (الركود) في حين يرتفع في بعض الفترات التوسعية بمعدل 3% أو 4% في كل ربع سنة.

تمرين رقم (9-1):

يحتوى الملف INCOME.XLS على بيانات متباطئات الدخل الشخصي والاستهلاك الشخصي.

(أ) احسب واشرح الإحصاءات الوصفية للدخل الشخصي والتغير في بيانات الدخل الشخصي.

(ب) ارسم واشرح أشكالاً مشابهة للشكلين (9-1 و 9-2) باستخدام متغير الاستهلاك الشخصي والتغير في الاستهلاك الشخصي.



شكل رقم (9-1): الدخل الشخصي في الولايات المتحدة

شكل رقم (9-2): التغير في الدخل الشخصي في الولايات المتحدة

هناك خاصية أخرى لبيانات السلاسل الزمنية لاتتوافر عادة في البيانات المقطعية، وهي وجود ارتباط على مستوى المشاهدات كافة. فعلى سبيل المثال، يرتبط الدخل الشخصي في الربع الحالي بصورة وثيقة بالدخل الشخصي في الربع السابق [3]. وفقاً للمفاهيم الواردة في الفصل الثامن، يرتبط متغير "الدخل الشخصي" بتغير "الدخل الشخصي لفترة تباطؤ واحدة". وفي الواقع، إذا قمنا بحساب الارتباط بين الدخل الشخصي والدخل الشخصي المتباطئ، فسوف تكون النتيجة 0.99971، وعند حساب الارتباط بين التغير في الدخل الشخصي والتغير في الدخل الشخصي المتباطئ، نحصل على -0.00235 وهذه النتائج تبدو بديهية. فالسلاسل الزمنية للمتغيرات الاقتصادية الكلية مثل: الدخل، إجمالي الناتج المحلي، الاستهلاك، ... الخ، تتغير ببطء شديد عبر الزمن. حتى في حالات الركود الشديد، لاتتراجع هذه المتغيرات بأكثر من 1% أو 2% في كل ربع. لذلك، يميل الدخل في هذا الربع إلى البقاء على المستوى ذاته تقريباً خلال الربع السابق وبدرجة ارتباط عالية بينهما. ومع ذلك، فإن التغير أو النمو في السلاسل الزمنية للمتغيرات الاقتصادية الكلية أو معدلات نموها يبدو غريباً. فالتغير في الدخل الشخصي خلال الربع الحالي والربع السابق قد يكونان مختلفين تماماً حيث يتضح ذلك من درجة الارتباط بينهما والتي تساوي الصفر تقريباً.

تم حساب الشكلين (9-1 و 9-2) ونتائج الارتباط الموضحة في الفقرة السابقة، باستخدام بيانات الدخل الشخصي في الولايات المتحدة. ولكن السلاسل الزمنية للمتغيرات الاقتصادية الكلية الأخرى في العديد من الدول الأخرى، تظهر سلوكاً مشابهاً. بعبارة أخرى يظهر Y اتجاهاً سلوكياً ذا درجة عالية من الارتباط عبر الزمن ولكن ΔY يظهر العكس، إذ إنه لايعكس أي اتجاه سلوكي فضلاً عن أن درجة الارتباط ليست عالية عبر الزمن. تعد هذه الخصائص في غاية الأهمية لنماذج الانحدار مع متغيرات السلاسل الزمنية لأنها ترتبط بصورة وثيقة بمشكلة

عدم الاستقرار؛ لذلك من الملائم تخصيص الجزء المتبقي من هذا الفصل لتطوير أدوات ونماذج منهجية للتعامل معها.

دالة الارتباط الذاتي:

يعطي الارتباط الذي أشرنا إليه أعلاه أمثلة مبسطة للارتباط الذاتي (Autocorrelation) (أي الارتباط الذي يتضمن متغيراً ومتباطئة للمتغير). وتعد دالة الارتباط الذاتي (Autocorrelation Function) أداةً عامة يستخدمها الباحثون لفهم خصائص السلاسل الزمنية. استناداً إلى الشرح في الجزء المعنون "لمحة عن المتغيرات المتباطئة" و "لمحة عن الرموز" الواردة في الفصل الثامن، سنستخدم تعبيرات مثل "الارتباط بين Y و Y المتباطئة" وسوف نرمز له بالرمز: r_1.

تمرين رقم (9-2):

يتضمن الملف INCOME.XLS بيانات عن الدخل الشخصي والاستهلاك في الولايات المتحدة.

(أ) لكل من هاتين السلسلتين على حدة، قم بإعداد رسم بياني بالمحورين X Y يوضح العلاقة بين المتغير والمتغير المتباطئ فترة واحدة.

(ب) احسب درجة الارتباط (r_1) بين هذين المتغيرين.

(ت) قم بأخذ الفروقات الأولى لهذين المتغيرين ثم كرر (أ) و (ب).

كيف يمكنك تفسير البيانات التي استنتجتها والارتباطات والرسم البياني.

بصفة عامة، ينصب اهتمامنا على الارتباط بين Y و Y المتباطئة لعدد p فترة زمنية. فعلى سبيل المثال، يتم رصد بيانات الدخل الشخصي على أساس ربع سنوي، ومن ثم فإن الارتباط بين Y و Y المتباطئة لعدد p فترة ($p=4$ فترات) يمثل الارتباط بين الدخل الحالي والدخل قبل سنة مضت (السنة تتكون من 4 أرباع). سنرمز لهذا الارتباط بالرمز r_ρ ونشير إليه على أنه (الارتباط الذاتي عند فترة التباطؤ p). تتعامل دالة الارتباط الذاتي مع r_ρ على أنه دالة في p (أي يتم حساب r_ρ حيث: $p=1,\dots\dots,p$). يمثل p الطول الأقصى للمتباطئة قيد الدراسة ويتم اختياره بحيث يكون كبيراً بدرجة مناسبة (أي $p=12$ بالنسبة للبيانات الشهرية). وتعد دالة الارتباط الذاتي إحدى الأدوات الأكثر استخداماً في تحليل السلاسل الزمنية الأحادية لأنها تظهر قدراً ملائماً من المعلومات عن تلك السلاسل.

لمحـــة:

(1) r_p هو الارتباط بين متغير (Y مثلاً) و Y متباطئ p فترة. لقد لاحظنا عند تناول r_1 أن Y_1 المتباطئ فترة واحدة هو Y_0 الذي لا وجود له. لذلك استخدمنا بيانات من t=2,…,T لتعريف Y المتباطئ وحساب r_1. وتواجه عملية حساب r_p مشكلة أكثر حدة. ضع في الاعتبار إمكانية استحداث متغير جديد (W) بمشاهدات تتمثل في $W_t=Y_t$ حيث: $t=p+1,\dots\dots,T$. يتمثل الارتباط بين W (أي Y) و Z (أي Y المتباطئة p فترة) في r_p. لاحظ أن كلاً من المتغيرين الجديدين يحتوي على $T-p$ مشاهدة. ومن ثم فإننا عند حساب r_p نستبعد ضمنياً المشاهدات الأولى البالغة p مشاهدة.

فإذا أخذنا في الاعتبار متباطئات ذات قيم عالية، فإننا نقوم بحساب الارتباط الذاتي لعدد قليل جداً من المشاهدات. وإذا اخترنا الحالة الأكثر تطرفاً (T= p) فلن تتبقى لنا أي مشاهدة يمكن استخدامها. وهذا هو السبب وراء عدم اختيار p ذي قيمة كبيرة جدا. تعد القضايا المطروحة في هذه الفقرة مشابهة لتلك التي طرحت في نماذج فترات التباطؤ الموزعة (انظر لمحة عن المتغيرات المتباطئة في الفصل الثامن).

(2) تتضمن دالة الارتباط الذاتي ارتباطات مع قيم متباطئات مختلفة. فمن الناحية النظرية، يمكننا استخدام بيانات من t=2,…,T لحساب قيمة r_1؛ بيانات من t=3,….,T لحساب قيمة r_2 ؛....الخ، انتهاء ببيانات من $t = p+1,......,T$ لحساب قيمة r_p. ولكن ينبغي ملاحظة أن ذلك يعني أن كل ارتباط ذاتي يتم حسابه بعدد مختلف من نقاط البيانات. ولهذا، فإن الممارسة التطبيقية هي اختيار المتباطئة ذات القيمة القصوى (p) واستخدام بيانات من t= p +1,…,T لحساب كافة الارتباطات الذاتية.

مثال: الدخل الشخصي في الولايات المتحدة:

يوضح الشكل رقم (9-1) دوال الارتباط الذاتي لـ: Y= الدخل الشخصي الأمريكي و ΔY = التغير في الدخل الشخصي (باستخدام بيانات من الملف INCOME.XLS) باستخدام متباطئة ذات قيمة قصوى مقدارها 12 ($p=12$).

يمكن أيضاً عرض هذه البيانات باستخدام الشكل البياني حيث يوضح المحور السيني قيمة المتباطئة والمحور الصادي درجة الارتباط الذاتي (الشكلان 9-3 و 9-4).

جدول رقم (9-1): دوال الارتباط الذاتي

قيمة المتباطئة (p) Lag Length (p)	الدخل الشخصي (Personal Income)	التغير في الدخل الشخصي (Change in Personal Income)
1	0.9997	-0.0100
2	0.9993	0.0121
3	0.9990	0.1341
4	0.9986	0.0082
5	0.9983	-0.1562
6	0.9980	0.0611
7	0.9978	-0.0350
8	0.9975	-0.0655
9	0.9974	0.0745
10	0.9972	0.1488
11	0.9969	0.0330
12	0.9966	0.0363

تتمثل إحدى الخصائص المميزة للجدول (9-1) والأشكال (9-3، 9-4) في أن قيمة الارتباطات الذاتية تميل إلى أن تكون من الناحية الافتراضية مساوية للواحد بالنسبة للدخل الشخصي حتى في حالة المتباطئات ذات القيم العالية. وعلى العكس من ذلك، تكون الارتباطات الذاتية للتغير في الدخل الشخصي صغيرة جداً وتظهر نمطاً عشوائياً؛ بعبارة أخرى تكون الارتباطات الذاتية مساوية للصفر. وهذا النمط يعد شائعاً في العديد من السلاسل الزمنية

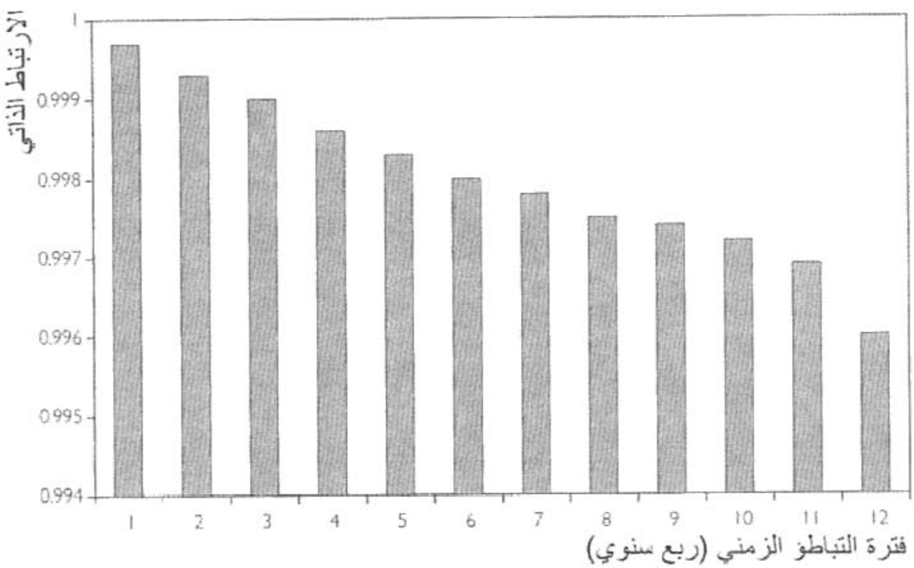
للمتغيرات الاقتصادية الكلية أو معظمها: حيث تكون الارتباطات الذاتية للسلسلة ذاتها قريبة من الواحد ولكن الارتباطات الذاتية للتغير في السلسلة تكون صغيرة جدا (أقرب للصفر). وفيما يلي عدد من وسائل التعبير عن هذه الارتباطات الذاتية:

(1) Y على درجة عالية من الارتباط عبر الزمن، حتى الدخل الشخصي قبل ثلاث سنوات (أي: $p=12$) على درجة عالية من الارتباط مع الدخل الحالي. ولكن ΔY لايظهر هذه الخاصية. فالنمو في الدخل الشخصي خلال هذا الربع لا ارتباط له مع النمو في الدخل خلال الأرباع السابقة.

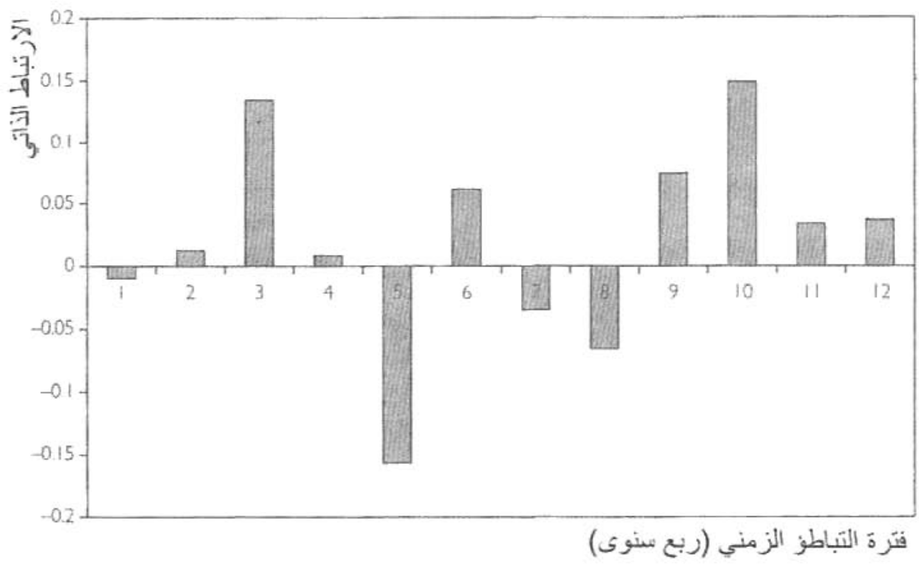
(2) إذا عرفت القيم السابقة للدخل الشخصي، فسوف تكون قادراً على إجراء تقدير جيد لحجم الدخل الشخصي في هذا الربع. ولكن معرفة القيم السابقة للتغير في الدخل الشخصي لن تساعدك في تقدير التغير في الدخل الشخصي خلال هذا الربع.

(3) بصفة عامة، Y يتضمن الماضي (أي أنه على درجة عالية من الارتباط بالقيم السابقة). وهذا مثال لسلوك التأثير المستمر وبعيد المدى (long memory). في حين لايتصف استخدام ΔY بهذه الخاصية.

(4) Y عبارة عن سلسلة غير مستقرة (nonstationary) في حين يعد ΔY سلسلة مستقرة (stationary). لم يتم تعريف مفهومي "غير مستقرة" و "مستقرة" منهجياً، ولكنهما مصطلحان مهمان في مجال الاقتصاد القياسي للسلاسل الزمنية. وسوف نتناولهما بتفصيل أكثر لاحقاً. ولكن يجب حالياً معرفة أن خصائص دالة الارتباط الذاتي للمتغير Y هي خصائص مميزة للسلاسل الزمنية غير المستقرة.



شكل رقم (9-3): دالة الارتباط الذاتي لبيانات الدخل



شكل رقم (9-4): دالة الارتباط الذاتي لبيانات التغير في الدخل

تمرين رقم (9-3):

استخدم بيانات (Y= الاستهلاك الشخصي) الموجودة في الملف INCOME.XLS:

(أ) احسب دالة الانحدار الذاتي لكل من Y و ΔY بمتباطئة ذات قيمة قصوى مقدارها 4 ($P=4$).

(ب) ضع هذه الدوال الخاصة بالارتباط الذاتي في شكل عمود بياني.

(ت) فسر النتائج المستخلصة من (أ) و (ب).

نموذج الارتباط الذاتي للسلاسل الزمنية الأحادية:

تعد دالة الارتباط الذاتي أداة مهمة لتحديد الصفات المميزة للسلاسل الزمنية. ومع ذلك، أوضحنا في الفصلين الثالث والرابع أن للارتباطات سلبياتها ومن ثم يمثل الانحدار أداة أكثر ملاءمة. وينطبق التبرير ذاته هنا، فالارتباطات الذاتية، بعبارة أخرى، هي مجرد ارتباطات، لذا قد يكون من الملائم إعداد نماذج أكثر تطوراً لتحليل العلاقات بين متغير ما ومتباطئاته.

لقد تمّ تطوير العديد من النماذج في الأدبيات الإحصائية والقياسية التي تتناول تحليل السلاسل الزمنية الأحادية، ولكن النموذج الأكثر استخداماً، والذي يمكن تفسيره أيضا كنموذج انحدار، هو "نموذج الانحدار الذاتي" (Autoregressive Model). وكما هو واضح من اسمه، فهو عبارة عن نموذج انحدار تكون فيه المتغيرات التفسيرية متباطئة للمتغير التابع (الانحدار الذاتي عبارة عن انحدار المتغير على متباطئاته الذاتية). ويتم اختصاره عادة بـ "AR".

سنبدأ بمناقشة نموذج الانحدار الذاتي مع المتغير التفسيري الذي يمثل المتغير التابع المتباطئ فترة واحدة؛ أي نموذج AR(1) (أو نموذج الانحدار الذاتي من الدرجة الأولى):

$$Y_t = \alpha + \phi Y_{t-1} + e_t$$

$$t=2,\ldots,T$$

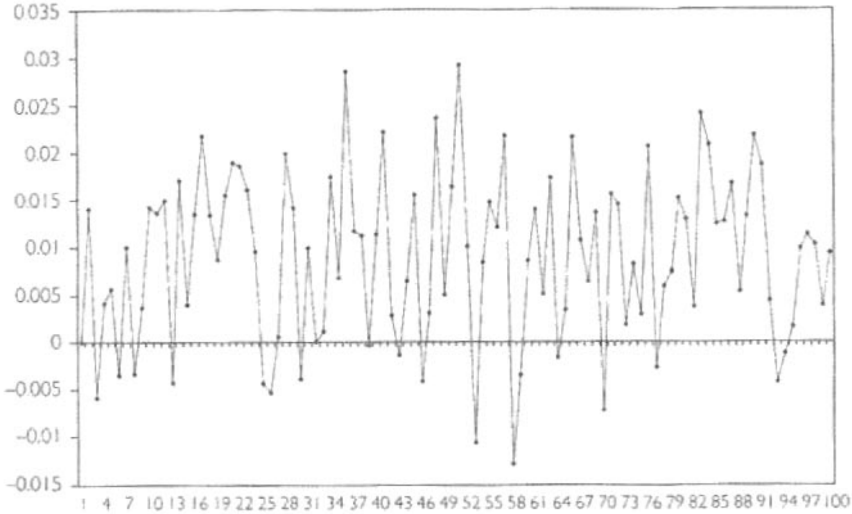
وهذا النموذج مشابه تماماً لنموذج الانحدار الذي تمت مناقشته في الفصول السابقة[4] باستثناء أن المتغير التابع هو Y_{t-1}. ترتبط قيمة ϕ في نموذج AR(1) ارتباطاً وثيقاً بسلوك دالة الارتباط الذاتي وبمفهوم عدم الاستقرار.

لفهم أنواع الخواص السلوكية لسلسلة AR(1)، دعنا نجري عملية محاكاة افتراضية لثلاث سلاسل زمنية مختلفة باستخدام ثلاثة خيارات مختلفة للمعامل ϕ:

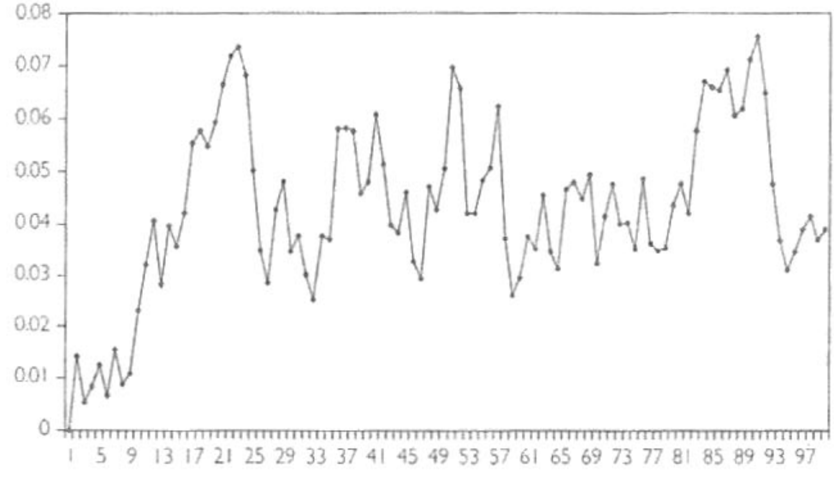
$$\phi = 1, \phi = 0.8, \phi = 0$$

وتعد قيمة α $(\alpha = 0.01)$ متساوية في السلاسل الثلاث وكذلك الأخطاء العشوائية (e). توضح الأشكال (9-5، 9-6 ، 9-7) عرضاً بيانياً للسلاسل الزمنية لمجموعات البيانات الثلاث.

لاحظ أن الشكل رقم (9-5) $(\phi = 0)$ يظهر تذبذبات (تقلبات) عشوائية حول متوسط مقداره نحو 0.01 (قيمة α). وفي الواقع فإنه مماثل تماماً للشكل رقم (9-2) الذي يتضمن عرضاً بيانيا للسلسلة الزمنية للتغير في الدخل الشخصي. ويوضح الشكل رقم (9-7) $(\phi = 1)$ سلوكاً اتجاهياً ويبدو مماثلاً للشكل رقم (9-1) الذي يوضح العرض البياني للدخل الشخصي. ويوضح الشكل رقم (9-6) $(\phi = 0.8)$ سلوكاً وسطياً بين التذبذبات العشوائية في الشكل رقم (9-5) والاتجاه العام القوي في الشكل رقم (9-7).

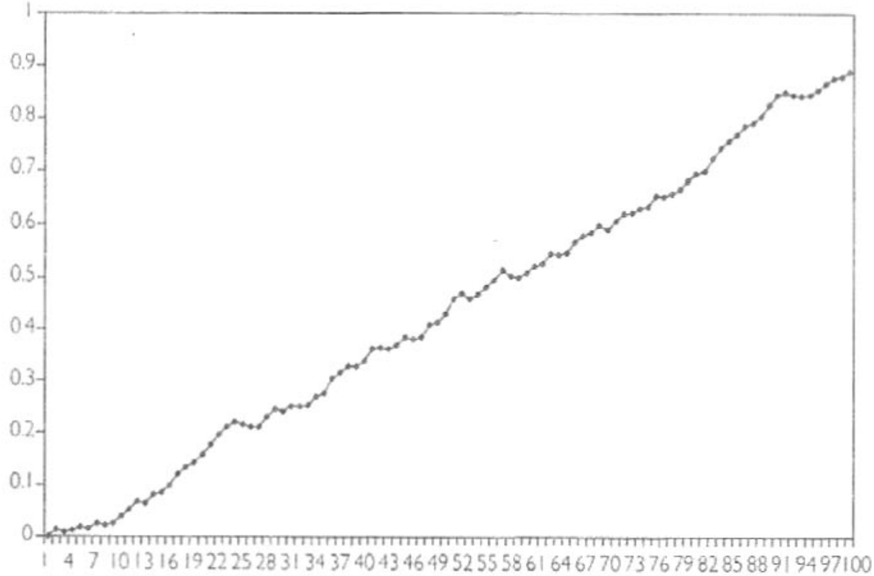


شكل رقم (9-5): السلسلة الزمنية لنموذج AR(1) عندما $\phi = 0$



شكل رقم (9-6): السلسلة الزمنية لنموذج AR(1) عندما $\phi = 0.8$

توضح الأشكال (9-5 - 9-7) أنواع السلوك التي قد تتبعها نماذج AR(1) وأسباب استخدامها الواسع في الاقتصاد الكلي. وعندما تختلف قيمة المعامل ϕ فستمكننا هذه النماذج من دراسة سلوك السلاسل الزمنية للاقتصاد الكلى التى تتسم بالتقلب مثل معدلات النمو، وكذلك دراسة سلاسل بيانات الاقتصاد الكلي التى تظهر في العادة اتجاهاً عاماً؛ أو للحالات التي تقع بين هاتين الحالتين الشاذتين.



شكل رقم (9-7): السلسلة الزمنية لنموذج AR(1) عندما $\phi = 1$

لاحظ أيضاً أن $\phi = 1$ يشير ضمناً إلى نوع السلوك الاتجاهى المشار إليه أعلاه بالسلوك الاتجاهى غير المستقر، في حين تشير القيم الأخرى إلى السلوك المستقر. وهذا يمكننا من إعداد تعريف اصطلاحي لمفهومي الاستقرار وعدم الاستقرار، على الأقل لنموذج AR(1): بالنسبة لنموذج AR(1)، يمكننا القول بأن Y مستقر إذا كان $|\phi| < 1$ وغير مستقر إذا كان $\phi = 1$.

ونادراً ما يتناول علم الاقتصاد الاحتمال الآخر المتمثل، $|\phi| > 1$. إن الاحتمال الأخير يعني ضمناً أن السلسلة الزمنية تظهر سلوكاً سريعاً (explosive) عبر الزمن. ولما كان مثل هذا السلوك غير الطبيعي (متزايد بشكل كبير جداً) تتم مشاهدته فقط في حالات غير عادية (التضخم الجامح مثلاً)، فإنه يعد ذا أهمية ضئيلة من الناحية العملية لذا لن نقوم بمناقشته هنا. ويبين الملحق رقم (9-1) التوضيح الرياضي لخصائص نموذج AR(1) وكيفية ارتباطه بقضية عدم الاستقرار.

تمرين رقم (9-4):

استخدم بيانات الملفات FIG95.XLS,FIG96.XLS,FIG97.XLS التي استخدمت في إعداد الأشكال (9-5) و (9-6) و (9-7) على التوالي.

(أ) احسب دالة الارتباط الذاتي لكل سلسلة زمنية باستخدام متباطئة بقيمة قصوى مقدارها 4.

(ب) اربط النتائج التي حصلت عليها في (أ) بنتائج التمرين رقم (9-3).

ركز بصفة خاصة على سؤال عما إذا كان نموذج AR(1) قادراً على محاكاة أنواع السلوك التى نلاحظها في بيانات السلاسل الزمنية لمتغير الاستهلاك الشخصي.

السلاسل الزمنية غير المستقرة والمستقرة:

لقد أدخلنا مصطلحي "غير مستقر" و "مستقر" دون تقديم أي تعريف منهجي (باستثناء نموذج AR(1)). وكما سنرى، فإن الفرق بين السلاسل الزمنية المستقرة وغير المستقرة في غاية الأهمية. إن التعريف المنهجي لهذين

المصطلحين يتطلب تناول قضايا إحصائية خارج نطاق هذا الكتاب، ولكننا سنقدم شرحاً عاماً لهذه المفاهيم أدناه.

من الناحية المنهجية، يركز الاقتصاديون عادة على نوع خاص من عدم الاستقرار يتواجد في العديد من السلاسل الزمنية الاقتصادية الكلية وهو عدم استقرار جذر الوحدة (Unit Root). سنعمل على تعميم هذا المفهوم لاحقاً. ولكن من المهم في هذه المرحلة اعتبار جذر الوحدة على أنه $\phi = 1$ في نموذج AR(1).

وفيما يلي نعرض الطرق المختلفة للتفكير بشأن ما إذا كان المتغير Y في السلسلة الزمنية مستقراً أو له جذر وحدة:

1- إذا كان $\phi = 1$ في نموذج AR(1)، فإن Y سيكون له جذر وحدة وإذا كان $|\phi| < 1$، سيكون Y مستقراً.

2- إذا كان لـ Y جذر وحدة، فإن ارتباطاته الذاتية ستكون قريبة من الواحد ولن تنخفض قيمه كثيراً مع زيادة قيمة المتباطئة.

3- إذا كان لـ Y جذر وحدة، فسوف يكون له تأثير مستمر (long memory). السلاسل الزمنية المستقرة ليس لها تأثير مستمر.

4- إذا كان لـ Y جذر وحدة، فإن السلسلة سوف تظهر سلوكاً اتجاهياً (خاصة إذا كان $\alpha \neq 0$).

5- إذا كان لـ Y جذر وحدة، فإن ΔY سيكون مستقراً. ولهذا السبب، يشار للسلسلة التي لها جذر وحدة غالباً على أنها سلسلة فروقات مستقرة (Difference Stationary).

ويمكن توضيح آخر نقطة بجلاء من خلال طرح Y_{t-1} من طرفي المعادلة في نموذج AR(1) للحصول على:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + e_t$$
$$(\rho = \phi - 1)$$

لاحظ أنه إذا كان $\phi = 1$ فإن $\rho = 0$ ويصبح بالإمكان كتابة المعادلة السابقة فقط في شكل ΔY_t مما يعني ضمناً أن ΔY_t تتقلب عشوائياً حول α. علاوة على ذلك، يجب ملاحظة أنه يمكننا اختبار $\rho = 0$ لمعرفة ما إذا كان للسلسلة جذر وحدة. وسوف تكون السلسلة الزمنية مستقرة إذا كان $-1 < \phi < 1$ ويعادل $-2 < \rho < 0$.

وسوف نشير إلى هذا بشرط الاستقرار (stationarity condition).

ولتقديم المزيد من الشرح لنموذج AR(1)، دعنا نتناول الحالة التي يكون فيها $\phi = 1$ (أو $\rho = 0$) و $\alpha = 0$.

وبالإمكان في هذه الحالة كتابة نموذج AR(1) على النحو التالي:

$$Y_t = Y_{t-1} + e_t$$

ويشار إلى هذا النموذج بنموذج المسار العشوائي (random walk). ولما كان $\phi = 1$ فإن Y سيكون له جذر وحدة وسيكون غير مستقر. ويعتقد أن هذا النموذج يصلح لظواهر مثل أسعار الأسهم. فالسعر الحالي للسهم هو سعره بالأمس زائدا حد خطأ (غير قابل للتوقع). فإذا لم تتبع أسعار الأسهم مساراً (تحركاً) عشوائياً، فإن التغير في سعر السهم يصبح قابلاً للتوقع ويصبح لدى المستثمرين إمكانات لتحقيق أرباح (arbitrage). ونظراً لوجود اعتقاد شائع بوجود عدد قليل جدا من مثل هذه الإمكانات الربحية المحتملة، فإن هناك قاعدة قوية للاعتقاد بأن أسعار الأصول المالية عديدة (أسعار الأسهم، أسعار الصرف،... الخ) يجب أن تتبع مساراً عشوائياً. وهذه الحجة تضفي قوة إضافية لفكرة وجود عدم استقرار في العديد من السلاسل الزمنية في الاقتصاد الكلى والتمويل.

مثال: الدخل الشخصي في الولايات المتحدة:

نموذج AR(1) عبارة عن نموذج انحدار ومن ثم يمكننا استخدام مربعات صغرى عادية لتقدير انحدار Y على Y المتباطئة [5]. فإذا قدرنا ذلك، فسوف نحصل على $\alpha = 0.039$ و $\phi = 0.996$. وبما أن تقدير المربعات الصغرى الاعتيادية، $\hat{\phi}$، والقيمة الحقيقية لمعامل AR(1)، ϕ، نادراً ما يتطابقان، فإنه من المحتمل جداً أن $\phi = 1$ نظراً لأن تقدير المربعات الصغرى الاعتيادية قريب جداً من الواحد. أما إذا قمنا بتقدير انحدار ΔY_t على Y_{t-1} فسوف نحصل على تقدير مربعات صغرى عادية β يساوي 0.004-.

لاحظ أننا نجد أن $\hat{\rho} = \hat{\phi} - 1$ كما كنا نتوقع تماماً.

تمرين رقم (9-5):

استخدم بيانات الملفات FIG95.XLS,FIG6.XLS,FIG97.XLS9 التي استخدمت في إعداد الأشكال (9-5)،(9-6)،(9-7) على التوالي.

(أ) احسب تقديرات المربعات الصغرى الاعتيادية ρ و ϕ في نوعي نموذج AR(1) المختلفين.

(ب) اربط النتائج التي حصلت عليها في (أ) بالسؤال الخاص بما إذا كانت تحتوي أي من السلسلتين على جذر وحدة.

(ت) كرر (أ) و (ب) باستخدام بيانات الاستهلاك الشخصي المتاحة في الملف INCOME.XLS

إضافات لنموذج الارتباط الذاتي من الدرجة الأولى AR(1):

لقد أوضحنا أعلاه أن نموذج AR(1) يمكن تفسيره كنموذج انحدار بسيط يمثل فيه Y للفترة السابقة المتغير التفسيري. ولكن بالإمكان إضافة المزيد من متباطئات Y كمتغيرات تفسيرية وذلك من خلال تعميم نموذج AR(1) إلى انحدار ذاتي من الدرجة p ، أي نموذج $AR(p)$:

$$Y_t = \alpha + \phi Y_{t-1} + \dots + \phi_p Y_{t-p} + e$$
$$(t = p+1, \dots, T)$$

لن نناقش خصائص هذا النموذج بأكثر من القول بأنها مماثلة لخصائص نموذج AR(1) وإن كانت أكثر عمومية في طبيعتها؛ بمعنى أن هذا النموذج قادر على محاكاة سلوك الاتجاه العام للسلاسل الزمنية الاقتصادية الكلية وسلوك التقلب العشوائي الذي تتسم به معدلات النمو.

عند بحث سلوك جذر الوحدة، من الملائم طرح Y_{t-1} من طرفي المعادلة السابقة. وبقدر من إعادة الترتيب[6] نحصل على الآتي:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots + \gamma_{p-1} \Delta Y_{t-p+1} + e_t$$

حيث إن معاملات هذا الانحدار $(\rho, \gamma_1, \dots, \gamma_{p-1})$ عبارة عن دوال بسيطة في $\phi_1, \dots\dots, \phi_p$. فعلى سبيل المثال: $\rho = \phi_1 + \dots + \phi_p - 1$. لاحظ أن ذلك مطابق لنموذج $AR(p)$ ولكنه مكتوب بطريقة مختلفة، ومن ثم سوف نشير للمعادلتين السابقتين بنماذج $AR(p)$. في حالة الشك في ما إذا كان الحد Y_{t-p} من المعادلة الأولى قد تحول إلى الثانية، لاحظ أنه يظهر في المعادلة الثانية في الحد ΔY_{t-p+1} $(\Delta Y_{t-p+1} = Y_{t-p+1} - Y_{t-p})$. لاحظ أيضاً أن كلتا المعادلتين لهما العدد

ذاته من المعاملات، $\rho+1$ (تتمثل معاملات الأولى في: $\alpha, \phi_1+...+\phi_p$ بينما في المعادلة الثانية ($\alpha, \rho, \gamma_1,..., \gamma_{p-1}$)، يتمثل المعامل الأخير لنموذج $AR(p)$ في γ_{p-1}. لاتدع ذلك يؤدى إلى إرباكك، فذلك ناتج فقط من إعادة ترتيب المعاملات في النموذج الأصلي.

إن النقاط الأساسية التي يجب ملاحظتها هنا هي أن المعادلة أعلاه لاتزال في شكل نموذج انحدار وأن $\rho = 0$ يعنى ضمناً أن السلسلة الزمنية Y لنموذج $AR(p)$ تحتوي على جذر وحدة، وإذا كان $-2 < \rho < 0$ ، فإن السلسلة تعد مستقرة. وبدراسة المعادلة السابقة مع $\rho = 0$ ، تظهر وسيلة مهمة للتفكر في سلسلة جذر الوحدة التي أشرنا إليها سابقاً: إذا كانت السلسلة الزمنية تحتوي على جذر وحدة، فإن نموذج الانحدار الذي يتضمن ΔY فقط هو الملائم (أي إذا كان $\rho = 0$ فإن الحد γ_{t-1} سوف يسقط من المعادلة وسوف تظهر فقط الحدود المتضمنة ΔY أو متباطئاته في الانحدار). ومن الشائع في هذا الصدد القول: "إذا وجد جذر وحدة، فبالإمكان أخذ فروقات البيانات للحصول على سلاسل بيانات مستقرة".

وكما سنناقش في الفصل التالي، مع استثناء حالة التكامل المشترك (Cointegration)، فإننا لانريد إدراج متغيرات جذور الوحدة في نماذج الانحدار. وهذا يعنى أنه في حالة وجود جذر وحدة في Y، فإننا سنعمل على أخذ الفروقات واستخدام ΔY. ولكي يتسنى لنا القيام بذلك، يجب علينا أولاً معرفة أن Y له جذر وحدة.

لقد أكدنا سابقاً أن السلاسل ذات جذور الوحدة تظهر اتجاهاً عاماً، فهل يعني ذلك أننا وبكل سهولة نستطيع اختبار الرسوم البيانية لـ Y فيما يختص بمثل هذا

الاتجاه لتحديد ما إذا كان هناك جذر وحدة؟ الإجابة هي لا. ولتوضيح السبب، دعنا نتناول نموذجاً آخر.

لقد أوضحنا آنفاً أن العديد من السلاسل الزمنية لمتغيرات الاقتصاد الكلي تتضمن اتجاهات عامة وأن نماذج AR مع جذور وحدة تحتوي ضمناً على سلوك اتجاهي. ومع ذلك، هناك نماذج أخرى تنطوي أيضاً على سلوك اتجاهي.

تخيل أن الشكل رقم (9-1) أو الشكل رقم (9-7) عبارة عن شكل انتشار للمتغيرين X,Y حيث يمثل المحور السيني (X) الزمن وأننا نريد بناء نموذج انحدار باستخدام هذه البيانات. ربما نحصل على خط الانحدار التالي:

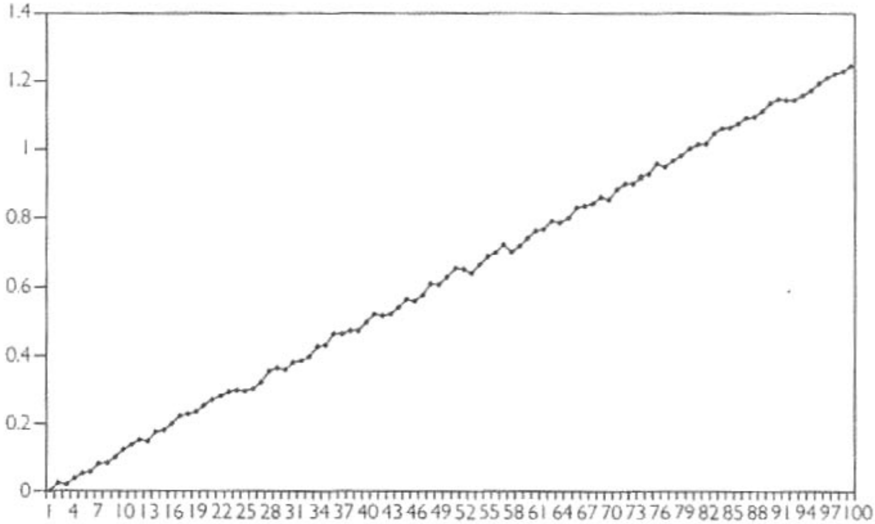
$$Y_t = \alpha + \delta t + e_t$$

حيث يمثل δ معامل المتغير التفسيري (الزمن) لتمييزه عن المعامل ϕ في نموذج AR(1). لاحظ أنه بإمكانك تفسير الانحدار السابق على أنه يحتوى على المتغير Y ومتغير أخر مع المشاهدات 1,2,3,4,...,T. وهذا نموذج انحدار آخر يؤدى إلى حدوث سلوك اتجاهي. يعرف δt بالاتجاه غير العشوائي (deterministic trend) لكونه دالة زمنية محددة. وعلى العكس من ذلك، تتضمن سلاسل جذور الوحدة ما يسمى "الاتجاه العشوائي" (stochastictrend) (يتضمن الملحق رقم (9-1) مبررات مصطلح الاتجاه العشوائي). بإمكاننا أيضاً دمج هذا النموذج في نموذج AR(1) للحصول على :

$$Y_t = \alpha + \phi Y_{t-1} + \delta t + e_t$$

الشكل رقم (9-8) وهو رسم بياني لسلسلة زمنية لبيانات افتراضية مستخلصة من النموذج السابق حيث $\delta = 0.01, \phi = 0.2, \alpha = 0$. لاحظ أن هذه السلسلة مستقرة نظراً لأن $|\phi| < 1$. ورغم ذلك، يبدو الشكل رقم (9-8) شبيهاً إلى حد

كبير بالشكل رقم (9-7) أو الشكل رقم (9-1). بإمكان النماذج المستقرة ذوات الاتجاه العام غير العشوائي أن تنتج رسومات بيانية لسلاسل زمنية شبيهة بتلك المستحدثة من النماذج غير المستقرة ذوات الاتجاه العشوائي. لذلك ينبغي تذكر أن النظر إلى الرسومات البيانية للسلاسل الزمنية وحدها ليس كافياً لتقرير ما إذا كان للسلسلة جذر وحدة.



شكل رقم (9-8): سلسلة زمنية مستقرة مع اتجاه عام

الطرح الوارد في الفقرة السابقة يحفزنا على مناقشة بعض المصطلحات التي سنستخدمها والتى يمكننا تلخيصها بالنقاط التالية:

1- متغيرات السلاسل الزمنية غير المستقرة التي سنركز عليها هي تلك التي تحتوي على جذر وحدة. فهذه السلاسل تتضمن اتجاهاً عاماً عشوائياً. ولكن إذا أخذنا الفروقات لهذه السلاسل الزمنية فإن السلاسل الزمنية المتبقية سوف تكون مستقرة. لهذا السبب، تسمى هذه السلاسل أيضاً سلاسل الفروقات المستقرة (difference stationary series).

2- السلاسل الزمنية المستقرة التي سنركز عليها يكون فيها $-2 < \rho < 0$ في نموذج $AR(p)$. ولكن هذه السلاسل قد تظهر سلوكاً اتجاهياً من خلال دمج اتجاه عام غير عشوائي. وفى هذه الحالة فإنها تسمي سلاسل مستقرة مع اتجاه عام (trend stationary).

تمرين رقم (9-6):

استخدمت بيانات الملف FIG98.XLS في إعداد الشكل رقم (9-8).

(أ) احسب دالة الارتباط الذاتي لهذه السلسلة المستقرة التي تتضمن اتجاهاً عاماً.

(ب) استناداً إلى إجابة البند (أ)، وضح ما إذا كانت دالة الارتباط الذاتي تمثل أداة مفيدة للتحقق من جذر الوحدة.

إذا أضفنا اتجاهاً محدداً (غير عشوائي) لنموذج $AR(p)$، سنحصل على نموذج عام مستخدم على نطاق واسع في تحليل السلاسل الزمنية الأحادية:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \dots\dots + \gamma_{p-1} \Delta Y_{t-p+1} + \delta t + e_t$$

وهذا النموذج يعرف بنموذج $AR(p)$ ذي الاتجاه غير العشوائي وسوف نستخدمه لاحقاً. وربما تندهش لعدم استخدامنا نموذج $AR(p)$ الأصلي الذي عرفناه في بداية هذا الجزء (والذي تتمثل متغيراته في $Y_{t-1}, \dots\dots, Y_{t-p}$).

هناك سببان لعدم استخدامنا هذا النموذج هما:

أولاً: أننا سنجري اختباراً للتحقق من جذر الوحدة. وفقاً للنموذج الحالي، يتعلق الاختبار بـ: $\rho = 0$. وفيما يختص باختبار عما إذا كانت المعاملات

صفرية، فقد عرفنا هذا الموضوع سابقاً (الفصل الخامس). ويعد الاختبار بغرض التحقق من جذر الوحدة باستخدام نموذج $AR(p)$ الأصلي أكثر تعقيداً.

ثانياً: غالباً ما تكون $Y_{t-1}, Y_{t-2}, ..., Y_{t-p}$ وثيقة الارتباط ببعضها (انظر دالة الارتباط الذاتي في الشكل رقم 9-3). فإذا أردنا استخدامها كمتغيرات تفسيرية في هذا الانحدار، فإننا حتماً سنواجه مشاكل تتعلق بالارتباط الخطي المتعدد (انظر الفصل السادس). ورغم ذلك، فإننا نستخدم في النموذج الحالي $Y_{t-1}, Y_{t-1}, ..., Y_{t-p+1}$ كمتغيرات تفسيرية ليست على درجة كبيرة من الارتباط (انظر الشكل رقم 9-4) ومن ثم نتجنب المشكلة.

مثال: الدخل الشخصي في الولايات المتحدة:

يحتوي الجدول رقم (9-2) مخرجات من انحدار مربعات صغرى عادية لـ: ΔY_t على $(Y_{t-1}, \Delta Y_{t-1}, \Delta Y_{t-2}, \Delta Y_{t-3})$ واتجاهاً زمنياً غير عشوائي. تم إعداد هذا الجدول باستخدام بيانات الدخل الشخصي المستخلصة من الملف INCOME.XLS. بعبارة أخرى، أنه يوفر مخرجات انحدار لنموذج AR(4) ذي الاتجاه غير العشوائي. إننا نشكك في أن هذه السلسلة من الدخل الشخصي تحتوي على جذر وحدة. وهذا الشك يعززه الجدول إلى حد ما. وبصفة خاصة، يوجد جذر الوحدة إذا كان ρ (المعامل الخاص بالمتغير التفسيري Y_{t-1}) مساوياً للصفر. وكما سنرى، فإن تقدير ρ صغير جداً $(\hat{p} = -0.018)$.

جدول رقم (9-2) نموذج AR(4) ذو الاتجاه المحدد

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
القاطع	0.138	0.108	1.279	0.203	-0.075	0.351
Y_{t-1}	-0.018	0.015	-1.190	0.236	-0.049	0.012
ΔY_{t-1}	-0.017	0.081	-0.217	0.829	-0.177	0.142
ΔY_{t-2}	0.014	0.081	0.172	0.863	-0.145	0.173
ΔY_{t-3}	0.130	0.080	1.627	0.106	-0.028	0.288
الوقت	0.00012	0.00012	0.955	0.341	-0.00013	0.00037

لاحظ أنه إذا تضمنت بيانات السلاسل الزمنية أنماطاً موسمية، فلا بد من إجراء المزيد من الإضافات على نموذج $AR(p)$ ذي الاتجاه غير العشوائي. وبالإمكان الحصول بسهولة على سلاسل زمنية تحتوي على أنماط موسمية. فعلى سبيل المثال، يشهد موسم عيد الميلاد مستويات عالية جداً من الإنفاق الاستهلاكي (أي في الربع الرابع من السنة). وبالمثل، يشهد فصل الصيف نسبة عالية من الإنشاءات في قطاع البناء والتشييد نظراً لملاءمة الطقس. وتظهر معدلات البطالة أيضاً تباينات موسمية حيث تكون في أدنى مستوياتها في الصيف عندما تكون معدلات التشغيل في قطاعي التشييد والزراعة في ذروتها. وتتسم كافة هذه المتغيرات بأنماط موسمية يمكن توقعها وقد نود إدراجها في نموذج الانحدار.

لاحظ أن الجهات الحكومية المعنية بالإحصاءات توفر بيانات خالية من الأنماط الموسمية (de-seasonalized). إن هذا النوع من البيانات يستبعد

الأنماط الموسمية ومن ثم لا داعي لأي مخاوف بشأنها. ولكن، في بعض الحالات، قد لاتتوافر مثل هذه البيانات أو قد تكون هناك حاجة للأنماط الموسمية في حد ذاتها [7]. ففي مثل هذه الحالات، ستكون في حاجة إلى معرفة كيفية التعامل مع هذا النوع الخاص من البيانات.

لايتناول هذا الكتاب الإجراءات اللازمة للتعامل مع الأنماط الموسمية بقدر من التفصيل. ولكن إحدى الطرق التي ينبغي معرفتها، تتضمن استخدام متغيرات صورية ويمكن تناولها هنا بإيجاز [8]. تذكر أن تحليل الانحدار يهدف في الأساس إلى وصف خواص المتغير التابع باستخدام المتغيرات التفسيرية. وبما أن الأنماط الموسمية قد تشكل خصائص مهمة للمتغير التابع، فإننا في حاجة إلى متغيرات تفسيرية لتوضيحها. هناك مجموعة متغيرات تفسيرية يمكن استخدامها في هذا الصدد وهى المتغيرات الصورية الموسمية (Seasonal Dummies). بإمكاننا إدراج مثل هذه المتغيرات الصورية كمتغيرات تفسيرية إضافية في نموذج $AR(p)$ ذي الاتجاه المحدد. فعلى سبيل المثال، يمكننا عند التعامل مع بيانات فصلية، استخدام المتغيرات الصورية التالية:

(1) $D_1 = 1$ إذا كانت المشاهدة تقع في الربع الأول ($D_1 = 0$ فيما عدا ذلك).

(2) $D_2 = 1$ إذا كانت المشاهدة تقع في الربع الثاني ($D_2 = 0$ فيما عدا ذلك).

(3) $D_3 = 1$ إذا كانت المشاهدة تقع في الربع الثالث ($D_3 = 0$ فيما عدا ذلك).

لاحظ أنك إذا أدرجت هذه المتغيرات الصورية الثلاثة في نموذج $AR(p)$ ذي الاتجاه غير العشوائي، لن تبرز أي مشكلات إحصائية جديدة. فالمربعات الصغرى الاعتيادية تعطي تقديرات جيدة للمعاملات كافة، وبإمكانك استخدام قيم الاحتمال (p) للتحقق ما إذا كانت المعاملات صفرية ... الخ.

بإيجاز فإنه يمكنك إضافة متغيرات صورية موسمية بصورة مباشرة إلى أي نموذج من نماذج السلاسل الزمنية التي تجدها في هذا الكتاب.

ربما تكون مندهشاً لإدراجنا المتغيرات الصورية لثلاثة مواسم فقط (D_1, D_2, D_3) في المثال السابق باستخدام البيانات الفصلية. إن سبب ذلك هو أننا لو كنا قد أدرجنا متغيراً صورياً رابعاً (D_4) للربع الرابع، لكنا قد واجهنا حالة ارتباط خطي متعدد تام (هذا التعبير يفترض وجود قاطع في النموذج). قد يكون من الصعب إلى حد ما توضيح السبب على أساس غير فني، ولكن إذا أردت التحقق من هذه العبارة بنفسك، أجر عملية الانحدار مع بيانات سلسلة زمنية موسمية باستخدام قاطع والمتغيرات الصورية (D_1, D_2, D_3, D_4) كمتغيرات تفسيرية وانظر ماذا يحدث. ولمزيد من الفهم لكيفية تفسير النتائج عند استخدام المتغيرات التفسيرية الصورية، عد مرة أخرى للفصل السابع، حيث يعد الجزء الذي يحمل عنوان: "الانحدار المتعدد مع وجود متغيرات تفسيرية صورية وغير صورية" وثيق الصلة بالموضوع.

اختبار الانحدار الذاتي من الدرجة (p) لنموذج الاتجاه العام غير العشوائي $AR(p)$:

لقد أوضحنا في الفصلين الخامس والسادس كيفية التحقق عما إذا كانت معاملات الانحدار صفرية. بالإمكان استخدام هذه الأساليب مع $AR(p)$ لنموذج الاتجاه العام غير العشوائي (أي إذا أردت حذف المتغيرات التفسيرية التي لاتختلف معاملاتها كثيراً عن الصفر). وبصفة خاصة، يتم الاختبار عادة للمساعدة في اختيار قيمة المتباطئة (p) وتحديد ما إذا كان للسلسلة جذر وحدة. وفي الواقع، فإنه من المعتاد إجراء الاختبار أولاً لاختيار قيمة المتباطئة وإجرائه بعد ذلك للتحقق من وجود جذر وحدة.

ورغم ذلك، هناك تحدٍ مهم يحدث في نموذج $AR(p)$ ولم يرد في الفصول الأولى. لفهم ذلك، دعنا نقسم معاملات النموذج إلى فئتين:

(1) $\alpha, \gamma_1, \ldots, \gamma_{p-1}$ و δ

(2) ρ.

بعبارة أخرى إننا ندرس اختبار فرضية تتضمن ρ بصفة مستقلة عن تلك التي تتضمن المعاملات الأخرى.

الاختبار لكل من $\alpha, \gamma_1, \ldots, \gamma_{p-1}$ و δ:

هناك العديد من المعايير الإحصائية وطرق الاختبار المتطورة التي تستخدم في تحديد قيمة المتباطئة الملائمة في نموذج $AR(p)$. ورغم ذلك فإن النظر إلى إحصاءات t أو قيم p في مخرجات الانحدار قد تكون مفيدة. فعلى سبيل المثال، فإن فحص بيانات الجدول رقم (9-2) أعلاه يوضح أن قيم p المرتبطة بالمعاملات على حدود ΔY المتباطئة غير معنوية وبالإمكان استبعادها من الانحدار (أي أن قيم p أكبر من 0.05). وكبديل لذلك، هناك مسار أكثر شيوعاً وهو العمل بالتتابع كما في نموذج فترات التباطؤ الموزعة؛ أي اختيار أقصى قيمة للمتباطئة ($p_{\max}$) والعمل تدريجياً على حذف قيم المتباطئة إذا كانت المعاملات وثيقة الصلة غير معنوية إحصائياً.

تحديداً ابدأ بـ $AR(p_{\max})$. فإذا كانت المتباطئة رقم $p_{\max}$ غير معنوية قم باختزال النموذج إلى $AR(p_{\max}-1)$ فإذا كانت المتباطئة $p_{\max}-1$ غير معنوية

في نموذج $AR(p_{max}-1)$ قم بحذفها واستخدم نموذج $AR(p_{max}-2)$ وهكذا. بصفة عامة، ينبغي البدء باختيار قيمة كبيرة لـ p_{max}.

في الارتباط الذاتي $AR(p)$ لنموذج الاتجاه العام غير العشوائي، يجب الاهتمام أيضاً باختبار $\delta=0$. ويمكن عمل ذلك بطريقة تقليدية من خلال التحقق عما إذا كانت قيمة p خاصة أقل من مستوى المعنوية (0.05 مثلاً). ويمكن إجراء هذا الاختبار في أي مرحلة ولكن من المعتاد إجراؤه بعد تنفيذ الإجراء التتابعي لاختيار p.

فيما يلي ملخص لطريقة هذا الاختبار:

الخطوة الأولى: اختر قيمة المتباطئة القصوى (p_{max}) التي تبدو ملائمة لك.

الخطوة الثانية: باستخدام المربعات الصغرى الاعتيادية، قدر $AR(p_{max})$ لنموذج الاتجاه العام غير العشوائي:

$$\Delta Y_t=\alpha+\rho Y_{t-1}+\gamma_1\Delta Y_{t-1}+......+\gamma_{P\max-1}\Delta Y_{t-P\max+1}+\delta t+e_t$$

فإذا كانت قيمة الاحتمال p لاختبار $\gamma_{P\max}-1=0$ أقل من مستوى المعنوية الذي اخترته (0.05 مثلاً)، اذهب إلى الخطوة الخامسة واستخدم p_{max} كقيمة للمتباطئة، وإلا فاذهب إلى الخطوة التالية.

الخطوة الثالثة: قدر نموذج $AR(p_{max}-1)$:

$$\Delta Y_t=\alpha+\rho Y_{t-1}+\gamma_1\Delta Y_{t-1}+......+\gamma_{P\max-2}\Delta Y_{t-P\max+2}+\delta t+e_t$$

فإذا كانت قيمة الاحتمال p لاختبار $\gamma_{P\max-2}=0$ أقل من مستوى المعنوية المختار (0.05 مثلا)، اذهب إلى الخطوة الخامسة. وإلا فاذهب إلى الخطوة التالية.

الخطوة الرابعة: قم بصورة متكررة بتقدير نماذج AR من الدرجات الدنيا لحين الحصول على نموذج $AR(p)$ الذي تكون فيه قيمة γ_{P-1} معنوية إحصائياً (أو ينفد عدد المتباطئات لديك).

الخطوة الخامسة: الآن أجر اختباراً للتحقق عما إذا كان الاتجاه العام غير العشوائي ينبغي حذفه، بمعنى أنه إذا كانت قيمة الاحتمال p لاختبار $\delta=0$ أكبر من مستوى المعنوية المختار، فاحذف متغير الاتجاه العام غير العشوائي.

مثال: الدخل الشخصي في الولايات المتحدة:

إذا طبقنا الطريقة السابقة على بيانات الدخل الشخصي بدءاً بـ $p_{\max}=4$، سيتم اختزال النموذج إلى:

$$\Delta Y_t=\alpha+\rho Y_{t-1}+e_t$$

أي أننا قدرنا أولاً AR(4) لنموذج الاتجاه العام غير العشوائي (انظر الجدول رقم 9-2) ووجدنا المعامل ΔY_{t-3} غير معنوي. لذلك، قدرنا AR(3) لنموذج الاتجاه العام غير العشوائي ووجدنا المعامل ΔY_{t-2} غير معنوي. ثم حذفنا المتغير الأخير وقدرنا AR(2)،... الخ. وأخيراً وبعد أن وجدنا أن المعامل غير معنوي، قمنا بتحديد النموذج المناسب ليكون AR(1). يتضمن الجدول رقم (9-3) نتائج تقدير المربعات الصغرى الاعتيادية لهذا النموذج.

جدول رقم (9-3): نموذج AR(1)

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
القاطع	0.039	0.014	2.682	0.008	0.010	0.067
Y_{t-1}	-0.004	0.002	-2.130	0.035	-0.0077	-0.0003

يوضح الجدول رقم (9-3) جدول النتائج التي قد تعرضها في ورقة أو مشروع بحثي مع وصف موجز ودقيق للطريقة التي استخدمتها للوصول إلى هذا النموذج النهائي.

تقودنا هذه النتائج إلى السؤال الأكثر أهمية بشأن الاختبار وهو: هل يحتوى Y على جذر وحدة؟ تذكر أنه إذا كان $\rho = 0$ فإن Y يحتوى على جذر وحدة. وفي هذه الحالة، يجب أخذ فروقات للسلسلة في نموذج الانحدار (أي أنها سلسلة فروقات مستقرة). وقد تعتقد أنه بالإمكان اختبار $\rho = 0$ بسهولة بنفس الكيفية التي اختبرت بها معنوية المعاملات الأخرى. فعلى سبيل المثال، فإن قيمة الاحتمال p للمعامل ρ في جداول إكسل هي 0.035. وهذه القيمة أقل من 0.05 ومن ثم فإنك قد تستنتج أن $\rho \neq 0$ وكذلك Y ليس له جذر وحدة. وهذا غير صحيح، ففي اختبار الفروض، يعد ρ مختلفاً عن باقي المعاملات، لذا ينبغي التعامل معه بصورة مختلفة.

الاختبار المتعلق بـ ρ:

للفهم الكامل للأسباب التي تحول دون إجراء اختبار جذر وحدة لفرضية $\rho = 0$ بالكيفية ذاتها التي اختبرنا بها المعاملات الأخرى، يجب أن تكون لديك معرفة بعلم الإحصاء تتجاوز محتوى هذا الكتاب. وتكفى الإشارة هنا إلى أن

معظم برامج الحاسب التي تستخدم في تقدير الانحدار، مثل إكسل، تفترض ضمنياً أن كافة متغيرات هذا النموذج مستقرة وذلك عند حساب قيم الاحتمال p . إذا كان المتغير التفسيري Y_{t-1} غير مستقر، فإن قيمة الاحتمال p المرتبطة به ستكون غير صحيحة. لقد تمكن اثنان من خبراء الإحصاء، ديكي وفولير، من تطوير أسلوب (طريقة) لاختبار جذر الوحدة. وهذه الطريقة تسمي "اختبار ديكي وفولير"[9]. ويوصي ديكي وفولير باستخدام إحصائية t لاختبار فرضية $\rho = 0$ ولكن مع تصحيح قيمة الاحتمال p .

بإمكاننا تحفيز استخدام اختبار ديكي وفولير، حيث أوضحنا في الفصل الخامس أن الاختبار يمكن إجراؤه من خلال مقارنة إحصائية الاختبار (إحصائية t في هذه الحالة) مع قيمة حرجة لتحديد ما إذا كان الأول إما "صغيراً" (وفى هذه الحالة يتم قبول الفرضية) أو "كبيراً" (وفى هذه الحالة يتم رفض الفرضية). ففي الحالة المعيارية (التى تكون فيها بيانات المتغير تحت الدراسة مستقرة)، يتم أخذ القيم الحرجة من جداول توزيع اختبار t. أوضح ديكي وفولير أن ذلك غير صحيح في حالة جذر الوحدة، وقاما بحساب الجداول الإحصائية الصحيحة لأخذ القيم الحرجة منها.

لقد كان الهدف من الفقرات السابقة هو توضيح أن إجراء الاختبار القياسي كان غير صحيح، وذلك لأنه لا يساعد كثيراً في إرشادك لما ينبغي فعله عملياً. فإذا كنت تود العمل بصورة شاملة مع بيانات السلاسل الزمنية، فمن الملائم:

(1) استخدام برنامج حاسب آلي أكثر ملاءمة لتحليل السلاسل الزمنية مقارنة ببرنامج إكسل. إن برامج مثل Microfit أو SHAZAM سوف تزودك تلقائياً بالقيم الحرجة الصحيحة أو قيم الاحتمال p لإجراء اختبار جذر الوحدة، وكما أوضحنا سابقاً، سيتم رفض فرضية جذر الوحدة إذا كانت قيمة

الاحتمال p أقل من 0.05 أو كان اختبار t أكبر من القيمة الحرجة (القيمة المطلقة).

(2) لا بد من قراءة المزيد في مجال الاقتصاد القياسي للسلاسل الزمنية ومعرفة كيفية استخدام جداول "ديكي وفولير" الإحصائية [10].

ورغم ذلك، يمكن استخدام قاعدة تقريبية: حيث لايسبب لك اتباعها خطأ كبيراً إذا كان عدد مشاهداتك كبيراً (T>50 مثلاً). وهذه القاعدة التقريبية تأتى ضمن الطريقة التالية لاختبار جذر الوحدة:

1- استخدم الطريقة الموضحة في الخطوات 1-5 أعلاه لتقدير $AR(p)$ لنموذج الاتجاه العام غير العشوائي. سجل إحصائية t المقابل لـ: p (أي، المعامل المتعلق بالمتغير Y_{t-1}).

2- إذا كانت النسخة النهائية من نموذجك تتضمن اتجاهاً عاماً غير عشوائي، ستكون القيمة الحرجة لاختبار ديكي وفولير -3.45 تقريباً. فإذا كانت إحصائية t للمعامل ρ أكبر سلبية من -3.45، أرفض فرضية جذر الوحدة وأخلص إلى أن السلسلة مستقرة. وإذا لم يكن كذلك، أخلص إلى أن للسلسلة جذر وحدة.

3- إذا لم تكن النسخة النهائية من نموذجك تتضمن اتجاهاً عاماً غير عشوائي، ستكون القيمة الحرجة لاختبار ديكي وفولير نحو -2.89. فإذا كانت إحصائية t للمعامل ρ أكبر سلبية من -2.89، أرفض فرضية جذر الوحدة وأخلص إلى أن السلسلة مستقرة. وإذا لم يكن كذلك، استنتج أن للسلسلة جذر وحدة [11].

إذا لم تكن النسخة النهائية من نموذج $AR(p)$ في المثال السابق تتضمن اتجاهاً عاماً غير عشوائي، وبلغت إحصائية t للمعامل p ‎-2.13 وهو لايعد أكثر سلبية من ‎-2.89 وبالتالي نقبل فرضية أن الدخل الشخصي يتضمن جذر وحدة. كن حذراً عند استخدام هذه القاعدة التقريبية عندما تكون إحصائية t مقاربة للقيم الحرجة الموضحة هنا.

تمرين رقم (9-7):

لقد أوصينا في هذا الفصل بطريقة تمكنك من البدء بارتباط ذاتي $AR(p)$ لنموذج الاتجاه العام غير العشوائي، اختيار قيمة متباطئة (p)، وتحديد ما إذا كان الاتجاه العام غير العشوائي ينبغي إدراجه أو استبعاده ومن ثم إجراء اختبار جذر وحدة. نفذ هذه الطريقة باستخدام السلسلة التالية:

(أ) السلاسل في الملف FIG95.XLS والملف FIG96.XLS (سلاسل مستقرة كما تعرف).

(ب)السلسلة في الملف FIG97.XLS (لها جذر وحدة كما تعرف).

(ت)السلسلة في الملف FIG98.XLS (لها اتجاه عام مستقر ولكنها تظهر سلوكاً اتجاهياً قوياً).

(ث)السلسلة الخاصة بالاستهلاك في الملف INCOME.XLS.

تمرين رقم (9-8):

قمنا في تمرين رقم (9-7) بإجراء اختبار للتحقق من جذور الوحدة في العديد من السلاسل. لقد لاحظنا أنه إذا كان للسلسلة جذر وحدة واحد، فإن سلسلة

فروقاتها سوف تكون مستقرة. تحقق من صحة ذلك لسلاسل لها جذور وحدة في تمرين رقم (9-7)؛ أي وضح كيف يمكنك إجراء اختبار للتحقق عما إذا كان التغير في السلسلة له جذر وحدة. بعد ذلك قم بإجراء هذا الاختبار.

بعض الملاحظات التحذيرية بشأن اختبار جذر الوحدة: يظهر اختبار ديكي وفولير مايطلق عليه خبراء الإحصاء اسم "الاختبار الضعيف" (Low Power). بعبارة أخرى، فإن الاختبار قد يقع في خطأ إيجاد جذر وحدة حتى في حالة عدم وجوده. من الناحية البديهية، قد تبدو سلسلة الاتجاه العام المستقرة مثل سلسلة جذر الوحدة (قارن الشكلين 9-7و9-8) ويكون من الصعب التمييز بينهما. بالإضافة إلى ذلك، فقد تظهر بعض الأنواع الأخرى لنماذج السلاسل الزمنية سلوك جذر الوحدة، في حين لا تحتوي هي في الحقيقة أي جذر وحدة. وخير مثال لذلك، نموذج السلسلة الزمنية التي تتميز بانقطاعات أو تغيرات فجائية. وهذه التغيرات الهيكلية (Structural Breaks) قد تحدث في نماذج السلاسل الزمنية للمتغيرات الاقتصادية الكلية نتيجة أحداث مثل الحروب أو الأزمات في الإمدادات (مثل الحظر النفطي بواسطة الأوبك). وقد تظهر التغيرات الهيكلية في أسعار الأسهم نتيجة لانهيار السوق، وفى أسعار السلع نتيجة للجفاف أو الكوارث الطبيعية الأخرى. خلاصة القول، أن التغيرات الهيكلية تمثل مشكلة للعديد من أنواع بيانات السلاسل الزمنية ولابد من أخذ الحيطة والحذر عند تفسير نتائج اختبارات ديكي وفولير.

ملخص الفصل:

1- تظهر العديد من السلاسل الزمنية سلوكاً اتجاهياً عاماً في حين لاتظهر فروقاتها سلوكاً مماثلاً.

2- تمثل دالة الارتباط الذاتي أداة عامة لتلخيص العلاقة بين متغير ما ومتباطئاته.

3- نماذج الانحدار الذاتي هي نماذج انحدار يتم استخدامها للتعامل مع متغيرات السلاسل الزمنية. ويمكن كتابة مثل هذه النماذج بطريقتين: الأولى مع Y_t كمتغير تابع والأخرى مع ΔY_t كمتغير تابع.

4- يعد التمييز بين النماذج المستقرة والنماذج غير المستقرة في غاية الأهمية.

5- السلاسل مع جذور الوحدة هي النوع الأكثر شيوعاً من أنواع السلاسل الزمنية غير المستقرة التي يتناولها علم الاقتصاد.

6- إذا كان لـ Y_t جذر وحدة، فبالإمكان تقدير نموذج $AR(p)$ مع ΔY_t كمتغير تابع وذلك باستخدام المربعات الصغرى الاعتيادية. تنطبق النتائج الإحصائية القياسية على كافة المعاملات عدا معامل Y_{t-1}.

7- اختبار ديكي وفولير هو اختبار للتحقق من وجود جذر وحدة، ويتضمن إجراء اختبار لمعرفة ما إذا كان معامل Y_{t-1} يساوي الصفر. إن برامج الحاسب الآلي، مثل إكسل، لاتعطي قيمة الاحتمال p الصحيحة للاختبار.

ملحق رقم (9-1): توضيح رياضي لنموذج الارتباط الذاتي من الدرجة الأولى AR(1):

بالإمكان فهم خصائص نموذج AR(1) رياضياً من خلال كتابة النموذج بطريقة مختلفة. ولأغراض التبسيط، نفترض أن $\alpha = 0$ للتركيز على الدور الذي تلعبه Y المتباطئة. لاحظ أن نموذج AR(1) يصلح لأي نقطة زمنية. لذلك يمكننا إبطاء معادلة AR(1) بأكملها، والتي وردت في هذا الفصل وكتابة ما يلي:

$$Y_{t-1} = \phi Y_{t-2} + e_{t-1}\cdot$$

وبتعويض هذه الصيغة الرياضية عن Y_{t-1} في نموذج AR(1) الأصلي، فإنه يمكن إعادة كتابة المعادلة على النحو التالي:

$$Y_t = \phi^2 Y_{t-2} + e_t + \phi e_{t-1}$$

لاحظ أن العبارة الجبرية السابقة تعتمد على Y_{t-2}، ولكن بالإمكان كتابة المعادلة على النحو التالي:

$$Y_{t-2} = \phi Y_{t-3} + e_{t-2}$$

وبتعويض هذه الصيغة الرياضية عن Y_{t-2} في المعادلة الأخرى. وبتكرار هذا الإجراء سنحصل على صيغة رياضية بديلة لنموذج AR(1):

$$Y_t = \phi^{t-1} Y_1 + \sum_{i=0}^{t-2} \phi^i e_{t-i}$$

وهذه العبارة تبدو معقدة ولكن يمكننا النظر في حالتين خاصتين كوسيلة لتفصيلها. سنقوم في الحالة الأولى بافتراض أن $\phi = 1$ ومن ثم يتم اختزال المعادلة السابقة إلى:

$$Y_t = Y_1 + \sum_{i=0}^{t-2} e_{t-i}$$

إن النقطة المهمة الجديرة بالملاحظة حول هذين الحدين في الطرف الأيمن من المعادلة السابقة هي أنهما يوضحان خاصية التأثير المستمر (Long Memory)، القيمة التي تبدأ بها السلسلة الزمنية عند Y_1، والتي تدخل دائماً في المعادلة الرياضية لـ Y_t حتى إذا كانت قيمة t كبيرة جداً. وهذا يعنى أن السلسلة الزمنية "لاتنسى أبدا" النقطة التي بدأت منها. كما أنها "لاتنسى" الأخطاء السابقة (مثلا، e_1 تدخل دائماً في المعادلة أعلاه عن Y_t حتى إذا بلغت قيمة t مقداراً كبيراً). ويمكن توضيح أن السلوك الاتجاهى لهذا النموذج ناتج عن الحد الثاني الذي يشير إلى أن Y الحالي يحتوي على مجموع كافة الأخطاء السابقة. يشير خبراء الإحصاء إلى هذه الأخطاء بالأخطاء "العشوائية"(stochastic) وإلى هذا النموذج عادة على أنه يتضمن اتجاهاً عاماً عشوائياً. وهذه خاصية أساسية للسلاسل غير المستقرة.

ثمة حالة خاصة ثانية وهي على النقيض من الخصائص الموضحة أعلاه. فإذا افترضنا أن $|\phi|<1$، يمكننا ملاحظة أن ϕ^{t-1} سيكون متناقصاً مع تزايد t (مثلا: إذا كان $\phi=0.5$، فإن $\phi^2=0.25$ و $\phi^{10}=0.001$ و $\phi^{100}=7.89\times10^{-31}$، ... الخ). وسوف يتناقص تأثير Y_1 والأخطاء السابقة على Y_t تدريجياً مع تزايد t ويبدأ Y في "نسيان الماضي" ببطء. ولن يظهر Y خاصية التأثير المستمر التي شاهدناها في حالة $\phi=1$. وهذه خاصية أساسية للسلاسل المستقرة.

ملاحظات ختامية:

1- تتوافر المعلومات التفصيلية عن اللوغاريثمات في أي كتاب عن الاقتصاد الرياضي. كما أنها قد تمت مناقشتها في الفصلين الثاني والرابع (انظر بصفة خاصة المناقشة الخاصة بالانحدار غير الخطي). تهدف هذه الملاحظة إلى تذكيرك بهذا الموضوع. فمن المعروف في الاقتصاد الكلي تناول اللوغاريثمات الطبيعية للسلسلة الزمنية إذا اتضح أنها متزايدة عبر الزمن. فإذا كانت السلسلة Y متزايدة بمعدل ثابت تقريبا، فإن السلسلة الزمنية للرسم البياني للوغاريثمات Y سوف تظهر خطاً مستقيماً. وفي هذه الحالة العامة، ستظهر لوغاريثمات Y أيضاً سلوكاً جيداً بصفة عامة. لاحظ أيضاً أنه في حالة انحدارات المتغيرات اللوغاريثمية، يمكننا تفسير المعاملات على أنها مرونات (Elasticities). بالإمكان توضيح أن لوغاريثم Y_t ناقصاً لوغاريثم Y_{t-1} أي $(\ln(Y_t)-\ln(Y_{t-1}))$ يساوي تقريباً النسبة المئوية للتغير في Y خلال الفترتين (t-1) و(t). لهذه الأسباب كافة، من الملائم عادة التعامل مع السلاسل اللوغاريثمية. لاحظ أن هذا التحويل اللوغاريثمي شائع جداً بحيث إن العديد من التقارير والبحوث توضح في البداية أن المتغيرات لوغاريثمية ولكنها بعد ذلك تغفل الإشارة الصريحة للتحويل اللوغاريثمي. فعلى سبيل المثال، قد يشير المؤلف إلى "اللوغاريثمات الطبيعية للدخل الشخصي" بـ "الدخل الشخصي" لغرض الاختصار. وسوف نتبع هذا التقليد في أمثلة هذا الكتاب.

2- بما أن Y_0 غير معروف، فإن ΔY_t يبدأ من t=2,......,T بدلاً عن t=1,......,T. سنركز على الحالة المفيدة عملياً وهي حالة الفرق الأول ولكن بالإمكان تعريف درجات أعلى للفروقات. فعلى سبيل المثال، فإن الفرق الثاني من Y_t يمكن تعريفه كالآتي:

$$\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}.$$

3- بطريقة أخرى، إذا عرفت قيمة الدخل الشخصي اليوم (1 تريليون دولار، مثلاً)، بإمكانك عمل تقدير جيد للقيمة المتوقعة في الربع التالي. وهذا يعني أنه يرتفع أو ينخفض بعدد من النقاط المئوية في شكل توسع أو انكماش. ولكن من غير المحتمل بدرجة كبيرة أن

يكون 500 بليون أو 1.5 تريليون دولار. إن هذه القدرة الجيدة للتوقع دليل للارتباط بدرجة عالية.

4- من الممارسات الشائعة، استخدام الحروف اليونانية للإشارة للمعاملات في نماذج الانحدار. بالطبع، يمكننا استخدام أي رمز يوناني نختاره للإشارة إلى معامل الميل في معادلة الانحدار. لقد اخترنا هنا ϕ بدلاً عن β الذي سنحتفظ به (ربما مع رمز سفلي) للإشارة للمعاملات المرتبطة بالمتغير التفسيري X.

5- تبرز بعض المشكلات الإحصائية غير المهمة مع تقدير المربعات الصغرى الاعتيادية في هذا النموذج، لاسيما إذا كان النموذج غير مستقر أو قريباً من ذلك (أي ρ مقارب جداً للصفر). ورغم ذلك، لاتزال المربعات الصغرى الاعتيادية وسيلة تقدير شائعة الاستخدام لنماذج AR مما يدل على أن هذه المشكلات ليست بذلك القدر من الأهمية. فإذا حضرت دورات دراسية في الاقتصاد القياسي أو إحصاءات السلاسل الزمنية مستقبلاً، فسوف تتعلم، حتماً، طرق تقدير أخرى.

6- تتضمن كل خطوة من خطوات اشتقاق هذه المعادلة المبادئ الرياضية البسيطة فقط (أي طرح الشيء ذاته من طرفي المعادلة،... الخ). ورغم ذلك، هناك العديد من الخطوات المستخدمة وأن اشتقاق هذه المعادلة ينطوي، في الحقيقة، على قدر كبير من العمليات الرياضية الشائكة.

7- ثمة سبب آخر حول لماذا تريد معرفة كيفية التعامل مع البيانات الموسمية وهو أن الإجراءات المتبعة بواسطة الجهات الحكومية لتنقيح البيانات من الأنماط الموسمية تصاحبها بعض المشكلات. ومن الصعب توضيح ماهية هذه المشكلات بمفهوم غير فني ومن ثم لن نناقشها في هذا الكتاب.

8- يخصص العديد من كتب السلاسل الزمنية جزءاً كبيراً لموضوع الموسمية. وخير مثال لذلك كتاب: Time Series Models for Business Economics and Forecasting, by Philip Hans Franses (Cambridge University Press,1998).

9- يستخدم بعض المؤلفين مصطلح "اختبار ديكي وفولير" لاختبار فرضية $p = 0$ في نموذج AR(1) بدلا عن مصطلح "اختبار ديكي وفولير الموسع" للاختبار في نموذج $AR(p)$ (أي يتم توسعة اختبار جذر الوحدة الأساسي بمتباطئات إضافية).

10- "الاقتصاد القياسي للمرحلة الجامعية"(Undergraduate Econometrics)، آر. كارتر هيل، وليام جريفيت وجورج جدج (الطبعة الثانية، 2000)، الفصل السادس يعد أفضل نقطة للبداية.

11- -3.45 و -2.89 هما القيمتان الحرجتان لـ: t=100 باستخدام مستوى معنوية مقداره 5%. تقع القيم الحرجة بالنسبة لـ: t بين 50 ولانهاية في إطار 0.05 من هذه القيم.

الفصل العاشــر

الانحدار ومتغيرات السلاسل الزمنية

(Regression with Time Series Variables)

يهتم الباحثون، عند إجراء وتحليل الانحدار، بقياس تأثير المتغيرات التفسيرية على المتغير التابع. وكما أوضحنا في الفصل الثامن، يكون هذا الهدف أكثر صعوبة عندما يستخدم الباحث بيانات السلاسل الزمنية نظراً لأن المتغير التفسيري قد يؤثر في المتغير التابع مع وجود فترة تباطؤ. وهذا يتطلب دائماً إضافة متباطئات المتغير التفسيري في الانحدار. بالإضافة إلى ذلك، وكما أوضحنا في الفصل التاسع، فإن المتغير التابع قد يكون مرتبطاً بمتباطئاته الذاتية مما يعنى أن متباطئات المتغير التابع أيضا يجب إضافتها في الانحدار.

هذه الاعتبارات تحفز نموذج الانحدار الذاتي مع وجود فترات تباطؤ موزعة (Autoregressive Distributed Lag (ADL) Model) المستخدم على نطاق واسع:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \dots\dots + \phi_p Y_{t-p} + \beta_0 X_t + \beta_1 X_{t-1} + \dots\dots\dots + \beta_q X_{t-q} + e_t.$$

في هذا النموذج، يعتمد المتغير التابع (Y) على p متباطئة ذاتية، القيمة الحالية للمتغير التفسيري (X)، و q متباطئة ذاتية للمتغير X. كما يسمح النموذج بوجود اتجاه عام غير عشوائي (t). وحيث يتضمن النموذج عدد p متباطئة للمتغير Y وعدد q متباطئة للمتغير X، فإننا سنشير له بالرمز $ADL(p,q)$ [1].

سنركز في هذا الفصل على حالة تتضمن متغيراً تفسيرياً واحداً هو X، ولكن بإمكاننا السماح بمتغيرات تفسيرية عديدة في التحليل.

يعتمد تقدير نموذج $ADL(p,q)$ وتفسيره على ما إذا كانت السلسلتان X و Y مستقرتين أم لا. سنتناول هاتين الحالتين كلاً على حدة. ورغم ذلك، ينبغي ملاحظة أننا نفترض باستمرار أن للسلسلتين Y و X خصائص الاستقرارية ذاتها؛ أي أن كلتيهما يجب أن تكون مستقرة أو أن لكلتيهما جذر وحدة. ومن المعروف،

بداهة، أن تحليل الانحدار يتضمن استخدام X لتفسير Y. فعلى سبيل المثال، من الصعوبة بمكان لسلسلة مستقرة تفسير تباين الاتجاه العام العشوائي في سلسلة جذر الوحدة. فذلك يعنى، من الناحية العملية، أنه يتعين عليك قبل إجراء انحدار السلسلة الزمنية، فحص الخصائص الأحادية للمتغيرات التي تود استخدامها. ويجب عليك، بصفة خاصة، إجراء اختبارات جذر الوحدة وفقاً للخطوات الموضحة في الفصل التاسع لكل متغير في التحليل.

انحدار السلاسل الزمنية عندما تكون كل من X و Y مستقرتين:

عندما تكون كل من X و Y مستقرتين، يمكن إجراء تقدير المربعات الصغرى الاعتيادية لنموذج الانحدار $ADL(p,q)$ بالطريقة القياسية الموضحة في الفصول(4-8). ويمكن إجراء اختبار معنوية المتغيرات باستخدام إحصائية t وقيمة الاحتمال p من برنامج حاسب آلي مثل إكسل. كما يمكن استخدام هذه الاختبارات بدورها في اختبار (p,q) (عدد متباطئات المتغير التابع والمتغير التفسيري على التوالي). ولكن يجب عليك ملاحظة أن التفسير الحرفي للنتائج سيكون مختلفاً إلى حد ما عن الحالة التطبيقية، كما هو موضح أدناه.

في حالة نموذج $AR(p)$ الوارد في الفصل التاسع، من الأفضل لنا، فيما يختص بتقدير المربعات الصغرى الاعتيادية وتفسير النتائج، إعادة كتابة النموذج باستخدام ΔY كمتغير تابع. وتصح الاعتبارات ذاتها بالنسبة لنموذج $ADL(p,q)$ الذي يمكن إعادة كتابته كما يلي:

$$\Delta Y_t = \alpha + \delta t + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \ldots + \gamma_{p-1} \Delta Y_{t-p+1}$$
$$+ \theta X_t + \omega_1 \Delta X_t + \ldots + \omega_q \Delta X_{t-q+1} + e_t.$$

يجب التأكيد على أن هذا النموذج هو نفس نموذج $ADL(p,q)$ الأصلي؛ وإن كان تعرض لبعض التغييرات الجبرية القليلة. وكما كان لدينا نوعان مختلفان من نموذج $AR(p)$ في الفصل التاسع، لدينا الآن نوعان مختلفان من نموذج $ADL(p,q)$. وكما فعلنا سابقاً، نستخدم أحرف يونانية جديدة لمعاملات الانحدار لتمييزها عن تلك المستخدمة في نموذج $ADL(p,q)$ الأصلي(2). وقد يبدو هذا النموذج معقداً، ولكنه لا يزال رغم ذلك مجرد نموذج انحدار. بمعنى أن هذا النموذج لا يتطلب أساليب رياضية جديدة لأنه قائم على معادلة خطية بسيطة.

لقد أوضحنا في الفصل التاسع أن السلاسل الزمنية للمتغيرات الاقتصادية الكلية غالباً ما تكون مرتبطة بصورة كبيرة بمتباطئاتها، وهذا يعني أن نموذج ADL يواجه مشكلة الارتباط الخطي المتعدد بصورة متكررة. ومن الناحية النموذجية، لن تواجهنا مثل تلك المشكلات مع النموذج في صيغته الجديدة. والأهم من ذلك، وكما سنرى، فإن هناك ميزة إضافية تتعلق بتفسير المعاملات. ولهذه الأسباب، سنتعامل مباشرة مع النسخة الثانية من نموذج $ADL(p,q)$.

لقد ناقشنا في الفصل السادس كيفية تفسير معاملات الانحدار مع التركيز بصفة خاصة على شروط ثبات العوامل الأخرى. تذكر أننا قد صغنا عبارات من شاكلة: "يقيس المعامل تأثير مساحة قطعة الأرض على سعر بيع المنزل مع بقاء العوامل الأخرى دون تغيير". ومازال بالإمكان استخدام مثل هذا التفسير في نموذج $ADL(p,q)$ ولكن ذلك ليس شائعاً على نطاق واسع. إذن، كيف يمكننا تفسير المعاملات في نموذج ADL؟ إن الطريقة الأكثر استخداماً تأتي عبر مفهوم المضاعف (multiplier). ربما تكون ملماً بهذه الفكرة نظراً لشيوعها على نطاق واسع في العلوم الاجتماعية؛ فعلى سبيل المثال، يستخدم الاقتصاديون المضاعف عند قياس تأثير التغير في الإنفاق الحكومي على الدخل الوطني. وهذا المفهوم يعد أكثر تعقيداً، في السياق الحالي، حيث ينبغي علينا تحديد توقيت للتأثير.

من الشائع التركيز على المدى البعيد أو المضاعف الكلي (total multiplier) وهذا ما سوف نفعله هنا. لتوضيح هذا المقياس، افترض أن X وY في وضع توازن أو ثبات؛ أي لا تتغيران عبر الزمن. ولكن فجأة تتغير X بمقدار وحدة واحدة مما يؤثر في Y التي تبدأ في التغير الأمر الذي يؤدي في نهاية المطاف وعلى المدى البعيد إلى قيمة توازنية جديدة. ويمكن تفسير الفرق بين قيمة Y التوازنية القديمة والجديدة بأنه تأثير X بعيد المدى على Y، وهو المضاعف بعيد المدى. يعد المضاعف على قدر كبير من الأهمية بالنسبة لراسمي السياسات الاقتصادية الذين يودون معرفة التأثيرات النهائية للتغيرات في تلك السياسات في المجالات المختلفة.

من الأهمية بمكان التأكيد على أن المضاعف بعيد المدى يقيس تأثير التغير الدائم في X. ففي الفقرة السابقة أخذت X قيمة معينة ثم تغيرت بصورة دائمة إلى مستوى جديد بمعدل وحدة واحدة عن القيمة الأصلية. يقيس المضاعف بعيد المدى تأثير مثل هذا التغير. وقد تهتم في بعض الحالات بتأثير تغير مؤقت في X (أي أن X تبدأ عند مستوى أصلي ثم ترتفع بوحدة واحدة لفترة واحدة قبل العودة إلى المستوى الأصلي مرة أخرى). لا يقيس المضاعف بعيد المدى تأثير هذا النوع من التغير. بإمكاننا استخدام تفسير التأثير الحدي (marginal effect) التقليدي لمعاملات الانحدار لمثل هذه التغيرات المؤقتة. إن المثال الوارد في الفصل الثامن، الذي ناقش تأثير التدريب في مجال السلامة على الخسائر الناجمة عن الحوادث، يوضح بعض طرق تحديد تأثير التغير المؤقت في المتغير التفسيري (مثلاً، لقد انصب اهتمامنا في ذلك المثال على تأثير الزيادة في التدريب في مجال السلامة خلال شهر معين على الخسائر الناجمة عن الحوادث. ولم نناقش تأثير زيادة التدريب في مجال السلامة بصفة دائمة).

وبالإمكان توضيح أن المضاعف بعيد المدى لنموذج $ADL(p,q)$ سيكون على النحو التالي (على الرغم من أننا لن نحاول إثبات ذلك هنا) [(3)]:

$$-\frac{\theta}{p}$$

بعبارة أخرى، إن المعاملات المتعلقة بالمتغيرين X_t و Y_{t-1} في نموذج ADL الجديد هي التي تعد مهمة للسلوك بعيد المدى، وهذا يعني أنه من السهولة بمكان الحصول على تقدير للمضاعف بعيد المدى.

من الضروري هنا التأكيد على أننا نفترض أن X و Y مستقرتان. لقد أوضحنا في الفصل التاسع كيف أن $\rho = 0$ في نموذج $AR(p)$ يعني ضمنياً وجود جذر وحدة. إن نموذج ADL ليس مماثلاً لنموذج AR. ولكن حتى يتسنى تقديم شرح تقريبي، لاحظ أنه إذا كان $\rho = 0$ ، فإن المضاعف بعيد المدى يكون لا نهائياً. وفي الواقع، يمكن توضيح أنه، ولكي يكون النموذج مستقراً، يجب أن يكون $\rho > 0$. [(4)] عملياً إذا كانت X و Y مستقرتين، سيتم استيفاء هذا الشرط.

مثال: تأثير مشتريات الحاسب الآلي في المبيعات:

درجت الشركات خلال العقد الماضي على شراء المزيد من أجهزة الحاسب الآلي على افتراض أن ذلك سيزيد إنتاجية العمل. إن الهدف من هذا المثال هو التحقق من هذا الافتراض عملياً. يتضمن ملف البيانات COMPUTER.XLS بيانات تم جمعها بواسطة شركة لمدة 98 شهراً حول مشترياتها من أجهزة الحاسب الآلي ومتغيراً يعكس إنتاجية القوى العاملة بالشركة في مجال المبيعات. وبصفة خاصة، فإن المتغير التابع والمتغير التفسيري هما:

Y= التغير النسبي في المبيعات بالمقارنة مع الشهر السابق.

X= التغير النسبي في مشتريات الحاسب الآلي بالمقارنة مع الشهر السابق.

يبلغ متوسط هذين المتغيرين 0.30% و 0.01% شهرياً مما يدل على أن هذه الشركة لم تزد إنفاقها على شراء أجهزة الحاسب الآلي في المتوسط بصورة كبيرة. ورغم ذلك، يجب ملاحظة أن هذا المتوسط يخفي تبايناً كبيراً. فقد ارتفع الإنفاق كثيراً في بعض الشهور في حين انخفض في شهور أخرى. وبافتراض أن كلا المتغيرين مستقر، يمكننا تقدير نموذج ADL(2,2) باستخدام المربعات الصغرى الاعتيادية. تذكر أنه إذا كانت متغيرات النموذج مستقرة، بالإمكان حساب كميات الانحدار القياسية (مثل تقديرات المربعات الصغرى الاعتيادية، قيم P، فترات الثقة) بالطريقة ذاتها الموضحة في الفصول أرقام (4-8). يتضمن الجدول رقم (10-1) نتائج هذه العملية.

باستخدام معادلة المضاعف بعيد المدى، سنحصل على تقدير مربعات صغرى اعتيادية مقداره: 1.042 =(0.125/-0.120)-. هناك طرق عديدة للتعبير عن هذه المعلومة فعلياً (تذكر أن المتغيرين التابع والتفسيري عبارة عن متغيرات نسبية):

1- ظلت مشتريات أجهزة الحاسب الآلي في هذه الشركة متزايدة في المتوسط بمعدل 0.01%في الشهر والمبيعات بمعدل 0.30% في الشهر. فإذا قررت الشركة زيادة ميزانية الحاسب الآلي بنسبة 1.01% في كل شهر (أي زيادتها بوحدة واحدة من 0.01 إلى 1.01)، فإن المبيعات على المدى البعيد سوف تبدأ في التزايد بنسبة 1.342% في الشهر (أي 0.30 الأصلي زائداً المضاعف بعيد المدى البالغ 1.042) [5].

2- يبلغ تأثير المضاعف بعيد المدى لمشتريات أجهزة الحاسب الآلي على المبيعات 1.042%.

3- إذا تزايد X بمعدل 1% بصفة دائمة، فإن قيمة Y التوازنية سوف تزيد بنسبة 1.042%.

جدول رقم (10-1): ADL(2,2) لنموذج الاتجاه العام العشوائي

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
قاطع	-0.028	0.041	-0.685	0.495	-0.110	0.054
Y_{t-1}	-0.120	0.013	-9.46	4.11E-15	-0.145	-0.095
ΔY_{t-1}	0.794	0.031	25.628	7.41E-43	0.733	0.856
X_t	0.125	0.048	2.605	0.011	0.030	0.221
ΔX_t	0.838	0.044	19.111	2.96E-33	0.750	0.925
ΔX_{t-1}	0.002	0.022	0.103	0.918	-0.041	0.046
الوقت	0.001	0.001	0.984	0.328	-0.001	0.002

ورغم ذلك، تشير المعلومات الإحصائية إلى أن ذلك قد لا يكون نموذجاً جيداً ما دامت بعض المتغيرات التفسيرية غير معنوية (مثلاً: قيم p للمعاملات الخاصة بالمتغيرات ΔX_{t-1} والاتجاه العام الزمني غير معنوية عند مستوى 5%). وهذا يثير مشكلة اختيار قيمة المتباطئة في نموذج $ADL(p,q)$. لن نناقش هذا الموضوع هنا بأكثر من القول بأن طريقة اختيار q في نموذج فترات التباطؤ الموزعة (انظر الفصل الثامن) وطريقة اختيار p في نموذج $AR(p)$ (انظر الفصل التاسع) يمكن دمجهما.

ليس هناك قاعدة عامة بشأن ما إذا كان ينبغي عليك اختيار p أولاً ثم اختيار q ثم تقرر عما إذا كان الاتجاه العام العشوائي ينبغي إدراجه أو يجب عليك عمل إجراء آخر (مثل اختيار q ثم p ثم الاتجاه العام أو اختيار q ثم الاتجاه العام ثم p، ... الخ). ومادمت حذراً فإن درجة الخطأ في اختيار نموذج لن تكون كبيرة.

تمرين رقم (10-1):

استخدم المتغيرات: Y= التغير النسبي في المبيعات و X= التغير النسبي في مشتريات الحاسب الآلي في ملف البيانات COMPUTER.XLS وحدد ما إذا كان النموذج المقدر في جدول رقم (10-1) يعد نموذجاً جيداً.

بالتحديد قم بما يلي:

أ- وضح ما إذا كان X و Y ليس لهما جذر وحدة (كما هو مفترض في المثال).

ب- ابدأ بنموذج ADL(3,3) مع اتجاه عام غير عشوائي، أجر اختبارات إحصائية لاختيار قيم المتباطئة المناسبة. هل تم تحقيق خيارات جيدة لكل من p,q في هذا المثال؟ هل كان يجب إدراج اتجاه عام عشوائي؟

ج- إذا وجدت أن المتغيرات ليست لها جذور وحدة وقمت بإجراء خيارات من p,q مختلفة عن خيارات المثال، احسب المضاعف بعيد المدى وقارنه مع نتيجة المثال.

تمرين رقم (10-2):

يحتوي ملف البيانات COMPUTER1.XLS متغيرات لها نفس شكل COMPUTER.XLS ولكن لشركة أخرى في قطاع مختلف (أو في صناعة مختلفة).

أ- أعد نفس التحليل المطلوب في تمرين رقم (10-1) باستخدام بيانات COMPUTER1.XLS; أي تحقق من أن كلاً من Y و X مستقرتان ومن ثم أجر اختباراً للحصول على نموذج $ADL(p,q)$ المناسب.

احسب المضاعف بعيد المدى للنموذج المقدر في (أ)

لمحة لمستخدمي برنامج إكسل (EXCEL):

أوضحنا في الفصل الثامن كيفية استحداث المتغيرات المتباطئة في إكسل باستخدام أوامر نسخ / لصق. بالإمكان استخدام أساليب مماثلة هنا. ومع ذلك، ينبغي ملاحظة أنه عند استحداث $\Delta X_t, \Delta Y_t$، فإنك سوف تستخدم معادلات. فإذا أردت مثلاً تعديل $\Delta X_t, \Delta Y_t$: لاستحداث $\Delta X_{t-1}, \Delta Y_{t-1}$، فيجب أن تكون حريصاً على نسخ ولصق القيم في الخلايا وليس المعادلات. ويمكنك عمل ذلك باستخدام خيار "لصق خاص" في إكسل.

انحدار السلاسل الزمنية عندما تكون كل من Y و X ذات جذور وحدة:

الانحدار الزائف:

فيما يلي من هذا الفصل، سنفترض أن Y و X ذات جذور وحدة. وبالطبع، ومن الناحية العملية، يتعين عليك إجراء اختبار للتحقق من ذلك باستخدام اختبار

"ديكي- فولير" في الفصل السابق. سنبدأ بالتركيز على حالة نماذج الانحدار بدون متباطئات ومن ثم البدء في إعداد نماذج مشابهة لنموذج $ADL(p,q)$:

افترض أننا مهتمون بتقدير الانحدار التالي:

$$Y_t = \alpha + \beta X_t + e_t$$

إذا كانت Y و X ذات جذور وحدة، فإن تقدير المربعات الصغرى الاعتيادية لهذا الانحدار قد يؤدي إلى نتائج خاطئة كلياً. فعلى سبيل المثال، إذا كانت قيمة β الحقيقية هي الصفر، فإن المربعات الصغرى الاعتيادية قد تعطي تقديراً ($\hat{\beta}$) يختلف تماماً عن الصفر. وقد تشير الاختبارات الإحصائية (باستخدام إحصائية t أو قيمة الاحتمال p) إلى أن $\beta \neq 0$. بالإضافة إلى ذلك، إذا كان $\beta = 0$، فإن R^2 يجب أن يساوي صفراً ($R^2 = 0$). ولكن في الواقع تكون قيمة R^2 غالباً كبيرة جداً.

بعبارة أخرى، إذا كانت Y و X ذات جذور وحدة، فإن كافة نتائج الانحدار الاعتيادية قد تكون مضللة وغير صحيحة. وهذه هي المشكلة المعروفة باسم "مشكلة الانحدار الزائف" (spurious regression). إننا لا نملك الأدوات الإحصائية لإثبات حدوث هذه المشكلة[6] ولكن من الأهمية بمكان معرفة الآثار السلبية لهذه المشكلة من منظور عملي. وباستثناء التكامل المشترك (Cointegration) الذي نشير إليه أدناه، يجب عدم إجراء انحدار Y على X إذا كانت للمتغيرات جذور وحدة.

انحدار السلاسل الزمنية عندما تكون كل من Y و X ذات جذور وحدة: التكامل المشترك:

يجب ألا يساورك أي قلق بشأن مشكلة الانحدار الزائف عندما تكون Y و X متكاملتين (cointegrated). فهذه الحالة لا تؤدي إلى التغلب على مشكلة

الانحدار الزائف فحسب، ولكنها توفر أيضاً قدراً من الفهم الاقتصادي. لقد حظي التكامل المشترك بقدر كبير من الاهتمام في الأبحاث الاقتصادية المعاصرة ومن ثم فإن هذا الموضوع يستحق البحث هنا بالتفصيل.

بإمكاننا الحصول على قدر من الفهم للتكامل المشترك بالنظر إلى الأخطاء في نموذج الانحدار أعلاه: $e_t = Y_t - \alpha - \beta X_t$.

بإعادة كتابة النموذج بهذه الكيفية، يتضح أن الأخطاء عبارة عن مزيج خطي لكل من Y و X. ورغم ذلك، تظهر كل من Y و X سلوك جذر وحدة غير مستقر بحيث قد تتوقع أن يظهر الخطأ أيضاً سلوكاً غير مستقر. (عموماً، إذا جمعت شيئين لهما خاصية مشتركة فإن النتيجة بصفة عامة تظهر تلك الخاصية)، فالخطأ عادةً له جذر وحدة. ومن الناحية الإحصائية، فإن جذر الوحدة هذا في حد الخطأ هو الذي يسبب مشكلة الانحدار الزائف. ولكن من المحتمل أن تعمل جذور الوحدة في Y و X على إلغاء بعضها ومن ثم يصبح الخطأ الناتج عن ذلك مستقراً. ففي هذه الحالة الخاصة التي تسمى "التكامل المشترك" (Cointegration)، تختفي مشكلة الانحدار الزائف ويصبح بالإمكان إجراء انحدار Y على X. خلاصة القول أنه إذا كانت Y و X ذات جذور وحدة ولكن كان مجموعهما الخطي مستقراً، فإنه يمكن القول أن Y و X متكاملتان[7].

إن الحدس وراء التكامل المشترك يكون أوضح في حالة $\alpha = 0$ و $\beta = 1$. تذكر ذلك عند قراءة العبارات التالية، وتذكر أيضاً أن المتغيرات ذوات جذور الوحدة تميل إلى إظهار سلوك اتجاه عام (مثلاً، إنها قد تكون متزايدة باطراد عبر الزمن وبالتالي قد تصبح كبيرة جداً).

1- إذا كانت لكل من Y و X جذور وحدة فسوف يكون لهما اتجاهات عشوائية. ولكن إذا كانتا متكاملتين، فلن يكون للخطأ اتجاه مماثل. وفي هذه الحالة لن يكون الخطأ

كبيراً وأن Y و X لن يبتعدا عن بعضهما؛ بعبارة أخرى، فإن Y و X سوف يتجهان معاً. وهذه الحقيقة تشجعنا على تناول المصطلحات الأخرى المستخدمة للإشارة إلى السلاسل الزمنية المتكاملة. فقد تسمع بالإشارة إليها على أن لها اتجاهات مشتركة. (common trends) أو (Co-trending)

2- إذا كنا نتحدث عن نموذج اقتصادي يتضمن مفهوم توازن، فإن e يمثل الخطأ التوازني. وإذا كانت Y و X متكاملتين، فإن الخطأ التوازني يظل صغيراً. ولكن إذا لم تكن Y و X متكاملتين، فإن الخطأ التوازني سيكون اتجاهياً وينحرف عن موضع التوازن ويتزايد بصورة كبيرة عبر الزمن. فإذا حدثت مثل هذه الانحرافات عن التوازن، فإن الكثيرين سوف يترددون عن القول بأن التوازن ذو مغزى.

3- إذا كانت Y و X متكاملتين، فإن هناك علاقة توازنية بينهما. وإذا لم تكونا متكاملتين، فلن توجد مثل هذا العلاقة (هذه مجرد إعادة صياغة لعبارة سابقة).

4- في الواقع الفعلي، من غير المحتمل أن يشهد النظام الاقتصادي توازناً دقيقاً نظراً لتعرضه للصدمات والتغيرات المفاجئة بشكل مستمر. من ناحية ثانية يجب أن لا تكون الانحرافات عن التوازن كبيرة جداً ولا بد أن يكون هناك دائماً اتجاه للعودة إلى حالة التوازن بعد حدوث صدمة. لذلك، إذا كان النموذج الاقتصادي الذي يحتوي ضمناً على علاقة توازنية بين Y و X صحيحاً، فسوف نلاحظ أن Y و X متكاملتان.

5- إذا كانت Y و X متكاملتين، فإن اتجاهاتهما سوف تلغي بعضها.

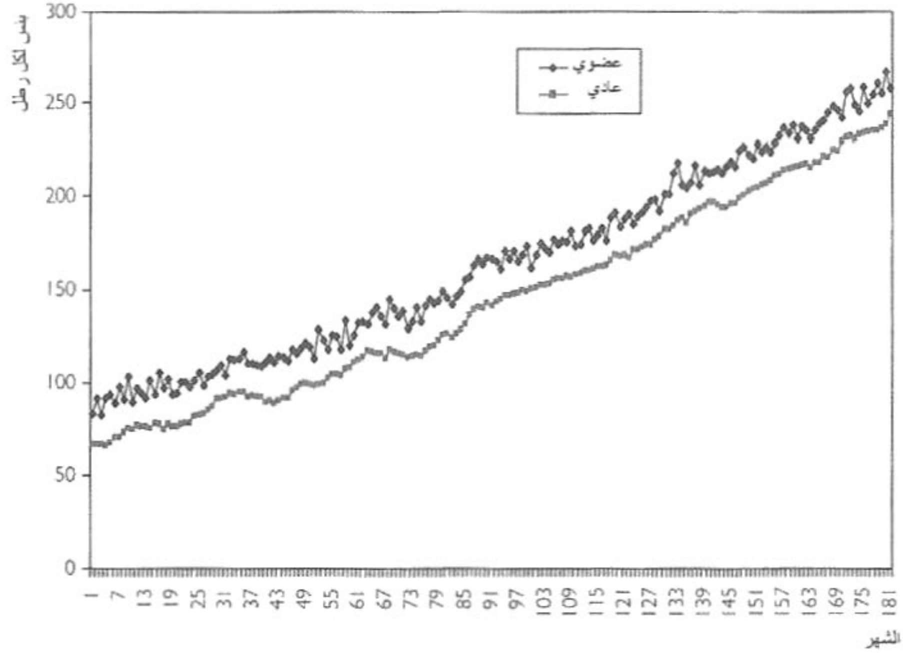
خلاصة القول: إذا كان التكامل المشترك قائماً، فإننا لانتجنب مشكلة الانحدار الزائف فحسب، بل نحصل أيضاً على معلومات اقتصادية مهمة (مثلاً: توجد علاقة توازنية أو أن سلسلتين تتجهان سوياً).

مثال: التكامل المشترك بين سعر سلعتين:

تشير النظرية الاقتصادية إلى أن السلع المتماثلة يجب أن تكون بدائل وثيقة لبعضها ومن ثم متكاملة. كمثال، يتضمن الملف ORANGE.XLS بيانات سلاسل زمنية لفترة 181 شهراً عن أسعار البرتقال العادي والبرتقال العضوي (يتم زراعته بدون استخدام أسمدة ومبيدات كيماوية) في سوق معينة. هاتان السلعتان على درجة كبيرة من الارتباط ولكن يرغب العديد من المستهلكين في دفع سعر أعلى قليلاً للبرتقال العضوي باعتبار أنه أفضل صحياً. إننا نتوقع تكاملاً مشتركاً بين سعر هاتين السلعتين نظراً لأن الفرق بين السعرين لا يمكن أن يزيد كثيراً. بمعنى أن العديد من الأفراد يرغبون في دفع زيادة طفيفة في السعر لشراء المنتجات العضوية. ولكن إذا أصبح الفرق كبيراً جداً فإنهم يتحولون إلى البرتقال العادي. فعلى سبيل المثال، قد يرغب العديد من المستهلكين في دفع 20 بنساً إضافياً لكل جنيه للحصول على الفوائد الصحية المزمعة ولكنهم ليسوا على استعداد لدفع 40 بنساً لكل جنيه. لذلك، إذا ارتفع سعر المنتجات العضوية مقارنة بالمنتجات الاعتيادية بصورة كبيرة، فإن العديد من الأفراد سيتوقفون عن شراء المنتجات العضوية ومن ثم فإن سعرها سوف ينخفض. ومن ناحية أخرى، إذا انخفض سعر البرتقال العضوي إلى نفس سعر البرتقال العادي تقريباً، فمن المحتمل أن يأكل القليل فقط البرتقال العادي، وفي هذه الحالة سوف ينخفض سعر البرتقال العادي.

خلاصة القول، أنه رغم أن أسعار هذين المنتجين سوف تتذبذب نتيجة لتقلبات العرض والطلب، إلا أن قوى السوق سوف تعمل دائماً على المحافظة على الفرق بين السعرين ثابتاً تقريباً. وهذا هو السر وراء مفهوم التكامل المشترك.

يوضح الشكل رقم (10-1) عرضاً بيانياً لأسعار هاتين السلسلتين ويقدم دليلاً قوياً مشاهداً بأن سعر هذين النوعين من البرتقال متكاملان. بمعنى أنه على الرغم من أن سعر البرتقال العضوي أعلى من سعر العادي، إلا أن سلوك الاتجاه العام للمتغيرين يبدو متشابهاً جداً.



شكل رقم (10-1) أسعار البرتقال العادي والعضوي

هناك العديد من الأمثلة الأخرى للتكامل المشترك، لاسيما في مجال الاقتصاد الكلي. إن أسعار الفائدة قصيرة وبعيدة المدى على سبيل المثال قد لا تتحرك سوياً بصورة دقيقة على المدى القريب ولكن من غير المحتمل انحرافها كثيراً على المدى البعيد. فإذا كانت أسعار الفائدة بعيدة المدى أعلى بكثير عن أسعار الفائدة

قصيرة المدى، فإن المتداولين سوف يشترون بالأسعار بعيدة المدى ويبيعون بالأسعار قصيرة المدى مما يؤدي إلى انخفاض الأولى وارتفاع الأخيرة وهذا المثال ينطوي على تكامل مشترك. هناك نظريتان اقتصاديتان أساسيتان تشيران ضمناً إلى وجود تكامل مشترك بين المتغيرات الاقتصادية الكلية، هما نظرية تعادل القوة الشرائية (purchasing power parity) وفرضية الدخل الدائم (permanent income hypothesis). وكذلك تم استخدام نظرية الطلب على النقود لتبرير نتائج التكامل المشترك. وبصفة عامة، يعد التكامل المشترك مفهوماً مهماً بالنسبة للمتخصصين في الاقتصاد الكلي.

التقدير والاختبار للمتغيرات المتكاملة:

لقد أوضحنا أعلاه، أنه إذا كان Y و X متكاملين، فإن مشكلة الانحدار الزائف لن تحدث، ومن ثم بإمكاننا إجراء انحدار المربعات الصغرى الاعتيادية Y على X والحصول على نتائج صحيحة. علاوة على ذلك فإن المعامل من هذا الانحدار هو المضاعف بعيد المدى. لذلك، ولأن الاهتمام منصب على المضاعف بعيد المدى، فإن التقدير مع المتغيرات المتكاملة سهل جداً.

قبل استخدام نتائج هذا الانحدار المعروف باسم الانحدار المتكامل (cointegrating regression)، من الأهمية بمكان التحقق من أن Y و X متكاملان. تذكر أنه إذا لم تكونا متكاملتين، فسوف تظهر مشكلة الانحدار الزائف وأن النتائج المتحصلة سوف تكون غير مجدية. إن فحص الرسم البياني لسلسلة زمنية، مثل الشكل رقم (10-1) قد يكون مفيداً ولكن تذكر أن الفحص البصري للأشكال البيانية يجب أن لا يكون بديلاً للاختبار الإحصائي.

هناك العديد من اختبارات التحقق من وجود تكامل مشترك فضلاً عن برامج الحاسب الآلي (مثل MicroFit) التي تمكنك من تنفيذ إجراءات معقدة جداً بلمسة

مفتاح. إن برنامج من نوع إكسل لا يمكنك من تنفيذ مثل هذه الاختبارات. ولحسن الحظ، فإن استخدام إمكانات الانحدار في برامج الجداول الإلكترونية، مع إجراء بعض المعالجات في البيانات، سيمكننا من إجراء اختبار واحد على الأقل للتحقق من التكامل المشترك.

يسمى اختبار التكامل المشترك الموضح هنا "اختبار أنجل - جرانجر" على اسمي خبيري الاقتصاد القياسي اللذين طورا هذا الاختبار. يستند هذا الاختبار إلى انحدار Y على X. تذكر أنه إذا حدث تكامل مشترك، فإن الخطأ الناجم عن هذا الانحدار سوف يكون مستقراً. وعلى العكس من ذلك، إذا لم يحدث تكامل مشترك، فسوف يكون للخطأ جذر وحدة. وفي ظل العلاقة الوثيقة بين الأخطاء والبواقي[8] سيكون من الملائم فحص خواص البواقي بهدف فحص التكامل المشترك الحالي. لقد ناقشنا في الفصل التاسع إجراء الاختبار للتحقق من وجود جذر وحدة في متغير السلسلة الزمنية. وسوف نقوم هنا بإجراء اختبار للتحقق من وجود جذر وحدة في البواقي وذلك باستخدام الأساليب ذاتها. وبصفة خاصة، يتضمن اختبار التكامل المشترك الخطوات التالية:

1- إجراء انحدار Y على X وحفظ البواقي.

2- إجراء اختبار جذر الوحدة على البواقي (دون إدراج اتجاه عام غير عشوائي).

3- في حالة رفض فرضية جذر الوحدة، ستكون Y و X متكاملتين. ولكن في حالة قبول فرضية جذر الوحدة لن يحدث أي تكامل مشترك.

لاحظ أنه بإمكانك، في برنامج إكسل، القيام بالخطوة الأولى بالنقر على الإطار المسمى "بواقي" في قائمة "انحدار" وتنفيذ التعليمات.

من الضروري هنا التأكيد على أن اختبار "إنجل - جرانجر" يستند إلى اختبار جذر الوحدة، ومن ثم ستبرز المشكلات الموضحة في الفصل التاسع. بعبارة

أخرى، وعلى الرغم من أن اختبار التكامل المشترك يستند إلى إحصائية t من نموذج انحدار (في هذه الحالة، انحدار يتضمن بواقي من انحدار أولي)، إلا أنه لا يمكنك استخدام قيم الاحتمال p المستخلصة من برنامج حاسب آلي غير متخصص، مثل إكسل. هناك جهات عديدة تنشر القيم الحرجة الصحيحة (التي تختلف بصورة طفيفة عن القيم الحرجة لاختبار ديكي- فولير). فإذا كنت تريد القيام بعمل كبير باستخدام بيانات السلاسل الزمنية، فمن الأفضل الاستفادة من الوقت في تعلم المزيد عن اختبار التكامل المشترك والبحث عن هذه الأرقام الحرجة الصحيحة. ومن ناحية أخرى، ولأغراض عديدة، من الملائم استخدام القاعدة التقريبية الموصى بها في الفصل التاسع.

لاحظ أنه عند اختبار جذر الوحدة في البواقي، لا ندرج اتجاهاً غير عشوائي. ففي حالة إدراج مثل هذا الاتجاه، سيعنى ذلك تزايد متوسط الأخطاء عبر الزمن، ومن شأن ذلك مخالفة فكرة التكامل المشترك (أي فكرة أن النظام الاقتصادي يعود دائماً إلى التوازن ومن ثم فإن الأخطاء لا تزداد بصورة كبيرة).

في ضوء هذه الاعتبارات، وعند إجراء اختبار جذر الوحدة على البواقي (الخطوة الثانية أعلاه)، استخدم 2.89- كقيمة حرجة لمقارنة إحصائية t معها. فإذا كان إحصائية t للمعامل ρ في انحدار جذر الوحدة الذي يتضمن بواقي أكثر سلبية من 2.89-، استنتج أن الأخطاء ليس لها جذر وحدة ومن ثم فإن Y و X متكاملتان.

تجدر الإشارة أيضاً إلى أنه سوف نقوم، في اختبار ديكي وفولير، باختبار فرضية أن $\rho = 0$ (أي أن فرض العدم هو جذر الوحدة). في اختبار التكامل المشترك نستخدم منهجية ديكي وفولير ولكننا نحصل على التكامل المشترك إذا رفضنا فرضية جذر الوحدة للبواقي. بعبارة أخرى، يتمثل فرض العدم في اختبار

إنجل – جرانجر في "عدم التكامل المشترك" ونستنتج أن "التكامل المشترك موجود" فقط عند رفض هذه الفرضية.

من الضروري أيضاً التأكيد على أنه نظراً لاستناد اختبار إنجل - جرانجر إلى اختبار ديكي - فولير، فإنه يواجه الصعوبات المذكورة في نهاية الفصل التاسع. أي أن اختبار إنجل - جرانجر أضعف وقد يكون مضللاً في حالة حدوث تغيرات هيكلية في البيانات.

مثال: التكامل المشترك بين سعر سلعتين:

دعنا نفترض أن لسلسلة سعر كل سلعة جذر وحدة. فإذا أجرينا انحدار Y= "سعر البرتقال العضوي" على X= "سعر البرتقال العادي" وذلك باستخدام بيانات من الملف ORANGE.XLS، فسوف نحصل على نموذج الانحدار التالي:

$$\hat{Y}_t = 20.686 + 0.996X_t.$$

تشير الطريقة أعلاه إلى أنه يجب علينا بعد ذلك اختبار جذر وحدة على البواقي، u_t، (يمكن إجراؤه باستخدام برنامج حاسب آلي مثل إكسل) من هذا الانحدار. إن الخطوة الأولى في عمل ذلك هي الاختيار الصحيح لقيمة المتباطئة باستخدام طريقة التتابع الموضحة في الفصل التاسع. افترض أننا عملنا ذلك وتوصلنا إلى أن نموذج AR(1) للبواقي مناسب. تشير طريقة ديكي - فولير إلى أنه يمكننا إجراء انحدار Δu_t على u_{t-1} ونحصل على النتائج الموضحة في جدول رقم (10-2).

إنجل – جرانجر في "عدم التكامل المشترك" ونستنتج أن "التكامل المشترك موجود" فقط عند رفض هذه الفرضية.

من الضروري أيضاً التأكيد على أنه نظراً لاستناد اختبار إنجل - جرانجر إلى اختبار ديكي - فولير، فإنه يواجه الصعوبات المذكورة في نهاية الفصل التاسع. أي أن اختبار إنجل - جرانجر أضعف وقد يكون مضللاً في حالة حدوث تغيرات هيكلية في البيانات.

مثال: التكامل المشترك بين سعر سلعتين:

دعنا نفترض أن لسلسلة سعر كل سلعة جذر وحدة. فإذا أجرينا انحدار Y= "سعر البرتقال العضوي" على X= "سعر البرتقال العادي" وذلك باستخدام بيانات من الملف ORANGE.XLS، فسوف نحصل على نموذج الانحدار التالي:

$$\hat{Y}_t = 20.686 + 0.996X_t.$$

تشير الطريقة أعلاه إلى أنه يجب علينا بعد ذلك اختبار جذر وحدة على البواقي، u_t، (يمكن إجراؤه باستخدام برنامج حاسب آلي مثل إكسل) من هذا الانحدار. إن الخطوة الأولى في عمل ذلك هي الاختيار الصحيح لقيمة المتباطئة باستخدام طريقة التتابع الموضحة في الفصل التاسع. افترض أننا عملنا ذلك وتوصلنا إلى أن نموذج AR(1) للبواقي مناسب. تشير طريقة ديكي - فولير إلى أنه يمكننا إجراء انحدار Δu_t على u_{t-1} ونحصل على النتائج الموضحة في جدول رقم (10-2).

جدول رقم (10-2): نموذج AR(1) باستخدام بواقي من انحدار تكامل مشترك

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
قاطع	0.024	0.292	0.083	0.934	-0.552	0.600
u_{t-1}	-1.085	0.075	-14.500	5.8E-32	-1.233	-0.938

تقول القاعدة التقريبية أن بإمكاننا مقارنة إحصائية t للمعامل u_{t-1} (الذي يبلغ -14.5) مع قيمة حرجة (-2.89). وبما أن القيمة الأولى أكثر سلبية من الأخيرة، فإننا نرفض فرضية جذر الوحدة ونخلص إلى أن البواقي ليس لها جذر وحدة. بعبارة أخرى، نستنتج أن السعرين متكاملان.

ولأننا قد وجدنا تكاملاً مشتركاً، فيجب أن لا تنزعج بشأن مشكلة الانحدار الزائف. عليه، بإمكاننا البدء في تفسير معاملاتنا دون أي خوف من أن تقديرات المربعات الصغرى الاعتيادية غير ذي معنى. إن حقيقة أن $\alpha = 20.69$ في الانحدار الأصلي Y على X يعكس زيادة مقدارها 20 بنساً لكل جنيه يرغب المستهلكون في دفعها للبرتقال العضوي بالمقارنة مع البرتقال العادي. بالإضافة إلى ذلك بلغت قيمة المضاعف بعيد المدى 0.996 مما يشير إلى أن ارتفاع سعر البرتقال العادي بمقدار بنس واحد سوف يؤدي إلى ارتفاع سعر البرتقال العضوي بمقدار 0.996 بنساً على المدى البعيد.

تمرين رقم (10-3):

استخدم بيانات الملف ORANGE.XLS في تكملة المثال السابق، وبصفة خاصة:

أ) أجر اختبار ديكي - فولير للتحقق من أن سلسلتي أسعار البرتقال لها جذور وحدة.

ب) أجر اختبار تتابع للتحقق من أن اختبار ديكي - فولير على البواقي قد تم بصورة صحيحة، أي هل نموذج AR(1) للبواقي مناسب؟

تمرين رقم (10-4):

يحتوي ملف إكسل LONGGDP.XLS بيانات سنوية عن نصيب الفرد من إجمالي الناتج المحلي الحقيقي لأربعة من أكبر الدول الناطقة بالإنجليزية (الولايات المتحدة، المملكة المتحدة، كندا وأستراليا) للفترة 1870- 1993 [9]. تحقق من أن هناك تغيرات أو اتجاهات عامة مشتركة بين إجمالي الناتج المحلي في هذه الدول. ويجب اتباع الخطوات التالية للإجابة عن هذا السؤال:

أ) إعداد رسم بياني للبيانات الموجودة في سلسلة زمنية ومناقشة النتائج (مثلاً، هل يتبع إجمالي الناتج المحلي اتجاهاً عاماً في جميع هذه الدول؟ هل يبدو أن هناك أنماطاً اتجاهية مشتركة بين هذه الدول؟).

ب) أجر اختبارات جذر وحدة على السلاسل الزمنية. ناقش النتائج التي ستحصل عليها.

جـ) بالنسبة للسلاسل الزمنية ذوات جذور الوحدة، أجر اختبارات تكامل مشترك. ابدأ بإجراء اختبار تكامل مشترك بين التشكيلات المختلفة المكونة من دولتين (مثلا، أولاً: الولايات المتحدة والمملكة المتحدة ثم الولايات المتحدة وكندا،... الخ). هل يبدو إجمالي الناتج المحلي متكاملاً في أي تشكيلة؟

د) لقد ذكرنا في هذا الفصل الخاص بالتكامل على الحالة التي تمت دراستها في الجزء (ج)، تحديداً في ظل وجود متغيرين فقط. باستخدام Y=USA والدول الأخرى كمتغيرات تفسيرية، أجر اختبار تكامل مشترك بين كافة السلاسل الزمنية. ناقش النتائج التي ستحصل عليها.

تمرين رقم (10-5):

استخدم بيانات Y= الدخل الشخصي وX= الاستهلاك الشخصي في ملف INCOME.XLS:

(أ) استخدم اختبار ديكي - فولير للتحقق من أن Y و X لهما جذور وحدة.

(ب) أجر انحدار Y على X واحفظ البواقي.

(ت) أجر اختبار جذر وحدة على البواقي باستخدام نموذج AR(1).

(ث) أجر اختبار جذر وحدة على البواقي باستخدام نموذج AR(2).

(ج) أجر اختبار جذر وحدة على البواقي باستخدام نموذج AR(3).

(ح) ماهو استنتاجك بشأن وجود تكامل مشترك بين Y و X؟ [10]

انحدار السلاسل الزمنية عندما تكون كل من Y و X متكاملتين: نموذج تصحيح الخطأ:

من المهم جداً، في التطبيقات العملية، إثبات أن Y و X متكاملتان. وكما أوضحنا أعلاه، يمكن ربط التكامل المشترك بفكرة الاتجاه المشترك لـ Y و X أو وجود علاقة توازنية بينهما. وتتمثل المهمة الأساسية الثانية في تقدير المضاعف بعيد المدى أو تأثير X على Y في المدى البعيد. وبالإمكان إجراء اختبار التكامل

المشترك وتقدير المضاعف بعيد المدى باستخدام انحدار Y على X. بناء على ذلك، وفي العديد من الدراسات التطبيقية لن تكون في حاجة أبداً لتجاوز هذا الانحدار البسيط، ولكن قد تكون مهتماً في بعض الأحيان بفهم السلوك قصير المدى على نحو قد يكون من المستحيل استخدام انحدار Y على X فقط. ففي مثل هذه الحالات يمكننا تقدير نموذج تصحيح الخطأ (Error Correction Model) أو باختصار (ECM).

تقول "نظرية جرانجر" أنه إذا كانت كل من Y و X متكاملتين، فإنه يمكن التعبير عن العلاقة بينهما من خلال نموذج تصحيح خطأ. سنفترض في هذا الجزء أن Y و X متكاملتان. ونماذج تصحيح الخطأ معروفة في الاقتصاد القياسي الخاص بالسلاسل الزمنية منذ عهد بعيد وأن نظرية جرانجر تعكس مدى شعبيتها. ولكي يتسنى لنا فهم خواص نماذج تصحيح الخطأ، دعنا نبدأ بهذا النموذج البسيط:

$$\Delta Y_t = \varphi + \lambda e_{t-1} + \omega_0 \Delta X_t + \mathrm{E}_t$$

حيث إن e_{t-1} هو الخطأ الذي تم الحصول عليه من نموذج انحدار Y على X (أي أن $e_{t-1} = Y_{t-1} - \alpha - \beta X_{t-1}$) وأن E_t هو الخطأ في نموذج تصحيح الخطأ. لاحظ أنه إذا عرفنا قيمة e_{t-1} فإن نموذج تصحيح الخطأ يصبح مجرد نموذج انحدار (على الرغم من استخدامنا بعض الحروف اليونانية الجديدة في هذا النموذج لتجنب حدوث اختلاط مع تلك المستخدمة في نماذج الانحدار الأخرى). وذلك يعني أن ΔY_t هو المتغير التابع و e_{t-1} و ΔX_t هما المتغيران التفسيريان. علاوة على ذلك، سنفترض أن $\lambda < 0$ [11].

لتفسير نموذج تصحيح الخطأ بشكل واضح، ضع في الاعتبار ضرورة أن يكون ΔY_t هو المتغير التابع. وكما أوضحنا في هذا الكتاب، يحاول نموذج الانحدار استخدام المتغيرات التفسيرية لتفسير المتغير التابع. ومع فهم هذه الحقيقة، لاحظ أن نموذج تصحيح الخطأ يوضح أن ΔY يعتمد على ΔX (أي أن التغير في X يحدث تغيراً في Y). بالإضافة إلى ذلك، يعتمد ΔY_t على e_{t-1}. وهذا الجانب الأخير يختص على نحو خاص بنموذج تصحيح الخطأ وهو الذي يضفي عليه هذا الاسم.

تذكر أن e يمكن اعتباره خطأ التوازن، فإذا كان $e \neq 0$، فإن النموذج سيكون غير متوازن. انظر الحالة التي يكون فيها $\Delta X = 0$ و e_{t-1} قيمة موجبة. وهذا الأخير يعني ضمناً أن Y_{t-1} كبيرة جدا بحيث لا يمكن أن تكون متوازنة (أي أن Y_{t-1} فوق مستوى توازن $(\alpha + \beta X_{t-1})$. وحيث إن $\lambda < 0$ فإن الحد λe_{t-1} سيكون سالباً وكذلك ΔY_t. بعبارة أخرى، إذا كانت Y_{t-1} فوق مستواها التوازني، فسوف تبدأ في التناقص في الفترة التالية وأن خطأ التوازن سيتم تصحيحه في النموذج، ومن هنا جاء مصطلح "نموذج تصحيح الخطأ" (12). وفي حالة أن يكون $e_{t-1} < 0$، فإن العكس هو الصحيح (أي أن Y_{t-1} ستكون دون مستواها التوازني ومن ثم سيكون $\lambda e_{t-1} > 0$ مما يجعل ΔY_t موجبة القيمة وبالتالي دفع Y للارتفاع في الفترة t).

بإيجاز، يتميز نموذج تصحيح الخطأ بخواص طويلة وقصيرة المدى. فالخواص الأولى (طويلة المدى) تشكل جزءاً لا يتجزأ من الحد e_{t-1} (تذكر أن β لا يزال هو المضاعف بعيد المدى وأن الأخطاء من انحدار يتضمن Y و X). أما السلوك قصير المدى فيتضح جزئياً بحد الخطأ التوازني الذي يوضح أنه إذا

كانت Y في وضع غير توازني، فسوف يتم جذبها نحو التوازن في الفترة التالية. ويتم توضيح المزيد من جوانب السلوك قصير المدى من خلال إدراج ΔX_t كمتغير تفسيري. وهذا الحد يعني أنه إذا تغيرت X فإن قيمة Y التوازنية سوف تتغير وأن Y ستتغير تبعاً لذلك. إجمالاً، يمكن القول بأن نموذج تصحيح الخطأ يتميز بخواص كثيرة ترتبط بصورة وثيقة بمفاهيم التوازن الاقتصادي.

يتميز نموذج تصحيح الخطأ أيضاً بخواص إحصائية جيدة، الأمر الذي يدعونا لعدم الانزعاج من مشكلة الانحدار الزائف. إن كلاً من Y و X لهما جذور وحدة ومن ثم فإن ΔY و ΔX مستقرتان. بالإضافة إلى ذلك، وبما أن X و Y متكاملتان، فإن الخطأ التوازني يعد مستقراً. لذلك يعد المتغير التابع وكافة المتغيرات التفسيرية في نموذج تصحيح الخطأ مستقرة. وهذه الخاصية تعني ضمناً إمكانية استخدامنا تقدير المربعات الصغرى الاعتيادية وإجراء الاختبار باستخدام إحصائية t وقيم الاحتمال p بالطريقة القياسية الموضحة في الفصل الخامس.

تبرز القضية الإحصائية الجديدة الوحيدة في نموذج تصحيح الخطأ نتيجة إدراج e_{t-1} كمتغير تفسيري. بالطبع، ليس بالإمكان مشاهدة الأخطاء في النموذج بصورة مباشرة. وهذا يثير مشكلة كيفية استخدامها كمتغير تفسيري في الانحدار. لقد تم استحداث أساليب اقتصاد قياسي متطورة لتقدير نموذج تصحيح الخطأ ولكن أسهل شيء يمكن عمله هو استبدال الأخطاء المجهولة ببواقي من انحدار Y على X (أي استبدال e_{t-1} بـ u_{t-1}). وهذا يعني أن الأسلوب البسيط المستند إلى انحدار مربعات صغرى اعتيادية يتم على النحو التالي:

الخطوة الأولى: أجر انحدار Y على X واحفظ البواقي.

الخطوة الثانية: أجر انحدار ΔY على ΔX والبواقي (من الخطوة الأولى) لمتباطئة فترة واحدة.

يجب التأكيد على أنه ينبغي، قبل القيام بهذا الإجراء التقديري ذي الخطوتين لنموذج تصحيح الخطأ، التحقق أن Yو X ذات جذور وحدة وأنهما متكاملتان.

لقد ناقشنا حتى الآن أبسط نموذج تصحيح خطأ، ولكن في الواقع العملي فإن نموذج تصحيح الخطأ له متباطئات تماماً كما لنموذج $ADL(p,q)$ متباطئات متغيرات تابعة وتفسيرية [13]. كما أن له اتجاهاً غير عشوائي. وبدمج هذه الخواص في نموذج تصحيح الخطأ، نحصل على:

$$\Delta Y_t = \varphi + \delta t + \lambda e_{t-1} + \gamma_t \Delta Y_{t-1} + \dots + \gamma_p \Delta Y_{t-p} + \omega_0 \Delta X_t + \dots + \omega_q \Delta X_{t-q} + \mathrm{E}$$

ما زال التعبير الرياضي في صيغة نموذج انحدار ويمكن تقديره باستخدام الإجراء ذي الخطوتين الموضح أعلاه. إن التكيف مع الوضع التوازني ينطبق أيضاً على هذا النموذج. فبالإمكان اتخاذ قرارات حول إمكانية إدراج اتجاه عام غير عشوائي وتحديد القيم الدقيقة لكل من p,q وذلك باستخدام إحصائية t وقيم الاحتمال p بنفس الكيفية التي تمت بها في نموذج ADL. وفي الحقيقة، يرتبط نموذج تصحيح الخطأ ارتباطاً وثيقاً بنموذج ADL إذ إنه يمثل نسخة محدودة منه.

مثال: التكامل المشترك بين سعر سلعتين:

لقد توصلنا في الجزء السابق من هذا المثال إلى أن: Y= سعر البرتقال العضوي وX= سعر البرتقال العادي، متكاملتان. وهذا يشير إلى إمكانية تقدير نموذج تصحيح خطأ. ولعمل ذلك، نبدأ بإجراء انحدار Y على X وحفظ البواقي (كما فعلنا في الجزء السابق من المثال). بعد ذلك يمكن إدراج البواقي (u_t) في الانحدار التالي (بصيغة المتباطئات):

$$\Delta Y_t = \varphi + \lambda u_{t-1} + \omega_0 \Delta X_t + \mathrm{E}_t$$

يوضح الجدول رقم (10-3) نتائج من تقدير المربعات الصغرى الاعتيادية لهذا النموذج:

جدول رقم (10-3): نموذج تصحيح خطأ بسيط

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
قاطع	-0.023	0.342	-0.068	0.946	-0.700	0.654
u_{t-1}	-1.085	0.075	-14.458	8.69E-32	-1.233	-0.937
ΔX_t	1.044	0.182	5.737	4.11E-08	0.685	1.403

بالإمكان تفسير المعلومات الإحصائية بالطريقة القياسية، إذ يمكننا القول بأن كافة المعاملات (عدا القاطع) معنوية من الناحية الإحصائية (ما دامت قيم الاحتمال p أقل من 0.05).

لقد لاحظنا من قبل أن $\hat{\beta} = 0.996$ ، وهذا هو تقدير المضاعف بعيد المدى. تلخص تقديرات λ و ω_0 في الجدول رقم (10-3) الخواص قصيرة المدى. ولتوضيح التفسير، لاحظ أن كافة متغيرات النموذج يتم قياسها بالبنسات. ويقيس معامل u_{t-1} (-1.085) مدى استجابة Y للأخطاء التوازنية. وبما أن المعامل ذو قيمة سالبة، فإن الأخطاء الموجبة تدفع ΔY بأن يكون سالباً ومن ثم انخفاض Y. ويعمل الخطأ التوازني البالغ بنساً واحداً، بصفة خاصة، على تخفيض Y بمقدار 1.085 بنساً في الفترة التالية، مع بقاء العوامل الأخرى دون تغيير. وهذا يعتبر استجابة سريعة جداً مع خطأ توازني. إن معامل ΔX_t =1.044 . بعبارة أخرى، تخيل ماذا يحدث إذا ظلت X دون تغير

لبعض الوقت ($\Delta X = 0$) ولكنها تغيرت فجأة بمقدار بنس واحد. يشير نموذج تصحيح الخطأ إلى أن Y سوف تتغير فوراً بمقدار 1.044 بنساً. وهذا يعني أن سعر البرتقال العضوي يستجيب بسرعة شديدة لتغيرات سعر البرتقال العادي. وقد تكون قابلية تعرض البرتقال للتلف هي التي تجعل البائعين يتفاعلون فوراً مع تغير سعر البرتقال العادي لضمان بيع البرتقال العضوي.

تمرين رقم (10-6):

استخدم بيانات ORANGE.XLS للتحقق من المثال السابق، مع الاهتمام الخاص بما إذا كان لنموذج تصحيح الخطأ عدد كاف من متباطئات ΔX و ΔY؟

تمرين رقم (10-7):

استخدم بيانات عن Y= الاستهلاك و X=الدخل الشخصي، من الملف INCOME.XLS. افترض (ربما بصورة غير صحيحة) أن Y و X متكاملتان.

أ) قدر نموذج تصحيح الخطأ. ابدأ بنموذج يحتوى على اتجاه عام عشوائي و $p = q = 4$ ثم أجر الاختبارات الإحصائية للحصول على نموذج تصحيح الخطأ الملائم.

ب) ناقش النتائج التي توصلت إليها، مع إعطاء اهتمام خاص لتقدير λ. ناقش ما يوضحه بشأن سرعة التكيف مع التوازن.

انحدار السلاسل الزمنية عندما تكون كل من Y و X ذوات جذور وحدة وليس بينهما تكامل مشترك:

من المحتمل مواجهة حالات تشير فيها اختبارات جذر الوحدة إلى أن السلسلة الزمنية ذات جذر وحدة، في حين يشير اختبار إنجل - جرانجر إلى أن السلسلة ليست متكاملة. بمعنى أن السلسلة لاتظهر اتجاهاً عاماً مشتركاً ولا تحتوي على علاقة توازنية. ففي مثل هذه الحالات، يجب عدم إجراء انحدار Y على X نظراً لمشكلة الانحدار الزائف. يشير وجود مثل هذه الخواص إلى ضرورة إعادة النظر في النموذج الأصلي واستخدام متغيرات تفسيرية أخرى. فبدلاً عن التعامل مع Y و X في ذاتهما، مثلاً، يمكن إجراء الفرق بينهما. (تذكر أنه إذا كان كل من Y و X ذات جذر وحدة، فإن ΔY و ΔX يجب أن يكونا مستقرين). وفي هذه الحالة، يمكنك التعامل مع التغيرات في سلسلتك الزمنية وتقدير نموذج ADL باستخدام الأساليب التي تم شرحها في بداية هذا الفصل. بعبارة أخرى، بإمكانك تقدير نموذج ADL الأصلي ولكن باستخدام التغيرات في المتغيرات:

$$\Delta Y_t = \alpha + \delta t + \phi_1 \Delta Y_{t-1} + \ldots + \phi_p \Delta Y_{t-p} + \beta_0 \Delta X_t + \ldots + \beta_q \Delta X_{t-q} + e_t$$

بالنسبة لمعظم متغيرات السلاسل الزمنية، يجب أن لا يعاني هذا النموذج من مشكلات الارتباط الخطي المتعدد. وكبديل لذلك، بإمكانك تقدير النسخة الثانية من نموذج ADL استناداً إلى بيانات الفروقات. ولكن إذا كنت تتعامل مع فروقات سلسلتك الزمنية ثم استخدمت نسخة نموذج ADL الذي يتضمن إجراء فروقات للبيانات، فسوف تحصل على بيانات الفروقات من الدرجة الثانية:

$$\Delta^2 Y_t = \alpha + \delta t + p\Delta Y_{t-1} + \gamma_1 \Delta^2 Y_{t-1} + \ldots + \gamma_{p-1} \Delta^2 Y_{t-p+1}$$
$$+ \theta\Delta X_t + \omega_1 \Delta^2 X_t + \ldots + \omega_q \Delta^2 X_{t-q+1} + e_t$$

حيث إن:

$$\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}$$

بالإمكان إجراء تقدير المربعات الصغرى الاعتيادية والاختبار بأي من هذين النموذجين بصورة مباشرة. وبصرف النظر عن النموذج المختار، يجب التأكيد على أن تفسير نتائج الانحدار سوف يتغير أيضاً.

تحديداً، دعنا نفترض أن Y= أسعار الصرف وX= أسعار الفائدة. فإذا كان كل من Y و X متكاملتين أو كانتا مستقرتين، بإمكاننا الحصول على تقدير التأثير بعيد المدى لتغير صغير في أسعار الفائدة على أسعار الصرف. أما إذا كانت Y و X غير مستقرتين ولا متكاملتين وأننا أردنا تقدير أي من المعادلتين السابقتين فإنه يمكننا الحصول على تقدير التأثير بعيد المدى لتغير صغير في التغير في أسعار الفائدة على التغير في أسعار الصرف. وقد يكون هذا، أو لا يكون، شيئاً محسوساً يمكن قياسه اعتماداً على الحالة قيد الدراسة.

لاحظ المثال الموضح في بداية الفصل عن تأثير مشتريات أجهزة الحاسب الآلي في المبيعات. في ذلك المثال، كانت المتغيرات في شكل نسب مئوية. فإذا كنا قد بدأنا بـ Y= المبيعات وX= مشتريات أجهزة الحاسب الآلي، لوجدنا أن لهما جذور وحدة ولكنهما غير متكاملتين. لذلك كنا سنواجه مشكلة الانحدار الزائف. وهذا هو السبب وراء استخدامنا للتغيرات النسبية.

تمرين رقم (10-8):

يتضمن ملف البيانات WP.XLS بيانات سنوية للفترة 1857 – 1987 عن X= الأجور، Y= الرقم القياسي لأسعار السلع الاستهلاكية في المملكة المتحدة [14].

هناك اعتقاد شائع بأن الضغوط بسبب رفع الأجور تعد سبباً رئيساً للتضخم، وأنك تريد التحقق من هذا الزعم بإجراء تحليل سلاسل زمنية لهذه البيانات، وخاصة:

أ) وضح بيانياً السلاسل الزمنية للأجور والأسعار. هل تظهر هذه السلاسل اتجاهات عامة مشتركة؟ هل يبدو أن هناك اتجاهاً عاماً مشتركاً؟

ب) أجر اختبار جذر وحدة لكل من X و Y. يجب إيجاد دليل على أن لكل منهما جذر وحدة.

ت) أجر اختبار تكامل مشترك لكل من X و Y. يجب تقديم دليل على أنهما غير متكاملتين.

ث) أجر فروقات للبيانات للحصول على ΔX و ΔY. كرر البندين (أ) و(ب) مع هذه المتغيرات الجديدة. يجب إثبات عدم وجود جذور وحدة.

حدد وقدر نموذج $ADL(p,q)$ باستخدام المتغيرات الجديدة ΔX و ΔY، وفسر نتائجك. لاحظ أن التغير في متباطئة مستوى السعر هو التضخم، أي أن ΔX و ΔY يمكن تفسيرهما بتضخم الأجور وتضخم الأسعار على التوالي.

ملخص الفصل:

1- إذا كانت جميع المتغيرات مستقرة، يمكن تقدير نموذج $ADL(p,q)$ باستخدام المربعات الصغرى الاعتيادية. تعد الأساليب الإحصائية متشابهة.

2- يتم غالباً استخدام نسخة مختلفة من نموذج ADL لتجنب مشكلة الارتباط الخطي المتعدد وعمل تقدير مباشر للمضاعف بعيد المدى.

3- إذا كانت جميع المتغيرات غير مستقرة، فيجب توخي الدقة والحذر في التحليل نظراً لمشكلة الانحدار الزائف.

4- إذا كانت جميع المتغيرات غير مستقرة ولكن خطأ الانحدار مستقر، فسوف يحدث تكامل مشترك.

5- لاتحدث مشكلة الانحدار الزائف في ظل وجود تكامل مشترك.

6- التكامل المشترك مصطلح يفضله الاقتصاديون لأنه يعنى ضمناً وجود علاقة توازنية.

7- بالإمكان إجراء اختبار التحقق من التكامل المشترك باستخدام اختبار إنجل- جرانجر. وهذا الاختبار عبارة عن اختبار ديكي - فولير على بواقي انحدار متغيرات متكاملة.

8- إذا كانت المتغيرات متكاملة فإنه يمكن استخدام نموذج تصحيح الخطأ. يوضح هذا النموذج السلوك قصير المدى على نحو لا يمكن الحصول عليه من خلال الانحدار المتكامل.

9- إذا كانت المتغيرات ذات جذر وحدة ولكنها غير متكاملة، فيجب عدم التعامل معها بصورة مباشرة. ولكن يجب أخذ فروقاتها ثم تقدير نموذج ADL باستخدام بيانات الفروقات. إن تفسير نتائج هذه النماذج قد يكون مضللاً.

ملاحظات ختامية:

1- من الناحية المنهجية، يمكن أن نسمي هذا نموذج $ADL(p,q)$ مع اتجاه عام غير عشوائي. ولكننا سنحذف العبارة الأخيرة للتسهيل. ففي الواقع العملي، ستلاحظ أن الاتجاه العام غير العشوائي غالباً غير معنوي وسوف يتم حذفه من النموذج بصورة أو أخرى. لاحظ أيضاً أن بعض الكتب المنهجية تختصر "فترات التباطؤ الموزعة للانحدار الذاتي" بالاختصار ARDL بدلاً عن ADL.

2- المعاملات التي تتضمن متباطئات المتغير التابع $p, \gamma_1, \ldots\ldots, \gamma_{p-1}$ هي ذاتها دوال $\phi_1, \ldots\ldots, \phi_p$ الموضحة في الفصل التاسع. كما أن $q+1$ معامل $(\theta, w_1, \ldots\ldots, w_q)$ مماثلة لدوال $\beta_0, \beta_1, \ldots\ldots, \beta_q$.

3- إن استخلاص المضاعف بعيد المدى من نموذج ADL ليس صعباً وبإمكانك محاولة ذلك في شكل تمرين. وهناك بعض العوامل المساعدة: افترض أن النموذج في وضع توازني لفترة طويلة وأن قيم X و Y التوازنية هي X^* و Y^* على التوالي. افترض الآن أن X قد ازدادت بصورة دائمة إلى X^*_{t+1}، وضح ماذا يحدث لـ Y.

4- كلمة "ثابت" عبارة عن مصطلح إحصائي لن يتم تعريفه في هذا الكتاب بصورة منهجية. ولكن يمكن تفسيره بصفة عامة كما يلي: إذا كان النموذج مستقراً، فإن ذلك يعني أن متغيرات السلسلة الزمنية لن تظهر اتجاهاً عاماً عشوائياً. وهذا المصطلح مماثل تقريباً لمصطلح "مستقر" (Stationary).

5- من الضروري التأكيد على أن تقدير المضاعف بعيد المدى هو 1.042 وبالإمكان حساب فترة ثقة ولكن الأمر سيتطلب اشتقاقات تعد خارج نطاق هذا الكتاب.

6- ربما تظن أن مشكلة الانحدار الزائف تحدث نتيجة تحيز متغير محذوف عندما يتم حذف المتباطئات من نموذج ADL، لكن الأمر أكثر من ذلك. حتى في حالة عدم وجود متباطئات في النموذج، تبرز مشكلة الانحدار الزائف.

7- لتوضيح كلمة "تكامل مشترك"، لاحظ أنه إذا كانت كل من X و Y ذات جذور وحدة فمن الشائع القول بأنهما متكاملتان. إن إضافة "co" للتأكيد على أن جذور الوحدة متشابهة أو مشتركة في X و Y تعطينا مصطلح تكامل مشترك (Cointegration).

8- تذكر أن الأخطاء عبارة عن اشتقاقات من خط الانحدار الحقيقي في حين أن البواقي هي اشتقاقات من خط الانحدار المقدر (انظر الفصل الرابع). إن رمز بواقي المربعات الصغرى الاعتيادية هو u_t.

9- لاحظ أن أياً من هذه السلاسل الزمنية تتكون من رقم قياسي (1913=100). فإذا تأملت البيانات، ستجد أن قيمة البيانات للمملكة المتحدة هي 64.85 عام 1870. إن حقيقة كون المتغيرات أرقاماً قياسية تعني أنه لايمكننا تفسير قيمة كل مشاهدة بالقول، مثلاً، أن نصيب الفرد من إجمالي الناتج المحلى في المملكة المتحدة كان 64.85 جنيهاً عام 1870. ومع ذلك، يمكننا تفسير التغيرات في السلسلة على أنها معدلات نمو إجمالي الناتج المحلى. وما هو أكثر أهمية بالنسبة لتحليل التكامل المشترك هو أن الاتجاه العام في الرقم القياسي لكل دولة يعكس بدقة تامة السلوك الاتجاهى لنصيب الفرد من إجمالي الناتج المحلى.

10- إذا أجبت عن هذا السؤال بصورة صحيحة ستجد أن التكامل المشترك موجود مع بعض قيم المتباطئات وغير موجود مع أخرى. وهذا أمر شائع في التطبيقات العملية، لذلك ينبغي ألا تنزعج منه. تشير النظرية الاقتصادية وبيانات الرسومات البيانية للسلاسل الزمنية بوضوح إلى أن التكامل المشترك يجب أن يحدث بين Y و X. ولكن اختبار انجل - جرانجر لا يشير إلى التكامل المشترك بصورة متسقة. ويمكن تفسير ذلك جزئياً بأن اختباري انجل - جرانجر وديكي - فولير ضعيفان.

11- لن نحاول إثبات لماذا ينبغي تحقق هذا الشرط سوى القول بأن شرط الاستقرار من النوع الذي تم بحثه في سياق نموذج $ADL(p,q)$.

12- هذا الشرح يوضح شرط الاستقرار $\lambda = 0$ الذي يؤكد أن الأخطاء التوازنية تم تصحيحها. فإذا كانت قيمة λ موجبة فإن الأخطاء التوازنية ستكون متضخمة.

13- لاحظ أننا لاندرج المزيد من متباطئات e_{t-1} كمتغيرات تفسيرية نظراً لانعكاسات "نظرية جرانجر"، التي لن نناقشها هنا.

14- هذه البيانات متباطئة ويعد هذا إجراء ضرورياً نظراً لأن بيانات الرقم القياسي لأسعار السلع الاستهلاكية على وجه الخصوص تبدو وكأنها تتزايد مع مرور الزمن.

الفصل الحادي عشر

تطبيقات لأساليب السلاسل الزمنية في الاقتصاد الكلي والتمويل

(Applications of Time Series Methods in Macroeconomics and Finance)

لقد ناقشنا في الفصول (8-10)، العديد من نماذج الانحدار المختلفة لمتغيرات السلاسل الزمنية. بالنسبة للعديد من الحالات، تعد معرفة هذه النماذج والأساليب المرتبطة بها (مثل اختبار التكامل المشترك) كافية للسماح لك بإعداد التقارير واكتساب الفهم الأساسي لخصائص البيانات. ولكن يستخدم الاقتصاديون في أغلب الأحيان نماذج وأدوات إحصائية أكثر تخصصاً. إن مناقشة بعضها في هذا الفصل تتيح فرصة لكسب المزيد من الخبرة في مجال تحليل السلاسل الزمنية. لذلك، سنقوم في هذا الفصل بمناقشة ثلاثة موضوعات مختلفة من السلاسل الزمنية التطبيقية: تقلبات (تذبذبات) أسعار الأصول، سببية جرانجر (Granger causality) ونماذج متجه الانحدار الذاتي (VAR). لقد لقي الموضوع الأول اهتماماً كبيراً في الآونة الأخيرة لاسيما من قِبل هؤلاء المهتمين بفهم خصائص الأصول (الأوراق) المالية مثل أسعار الأسهم، أما الآخران فهما على قدر كبير من الأهمية في مجال الاقتصاد الكلي.

تقلبات أسعار الأصول:

سنناقش في هذا الجزء نموذج انحدار بسيط يعد ذا أهمية لفهم تقلبات أسعار الأصول (أسعار الأسهم مثلاً). يرتبط هذا النموذج ارتباطاً وثيقاً بنماذج الانحدار الذاتي المشروط بعدم ثبات التباين (Autoregressive Conditional Heteroskedasticity) (ARCH) وعلى الرغم من هذا الاسم المخيف، إلا أن نماذج ARCH تعد أداة تحظى باهتمام متزايد وسهلة الاستخدام نسبياً في مجال المالية. إن البحث الكامل لنماذج ARCH يعد خارج نطاق هذا الكتاب. ولحسن الحظ، لن يكون من الصعب عليك استيعاب المفاهيم والأفكار الأساسية لهذه النماذج وتطوير أدوات انحدار بسيط للتعامل معها. سنركز في هذا الجزء على حالة التقلبات في أسعار الأصول (مثل أسعار الأسهم، أسعار الصرف، ... الخ). تجدر

الإشارة في هذا السياق إلى أن مفهوم التقلبات يلعب دوراً مهماً في مجالات الاقتصاد الأخرى مثل الاقتصاد الكلي.

لتقديم بعض الشرح، تذكر مناقشتنا لنموذج المسار (التحرك) العشوائي في الفصل التاسع، إذ عرفنا النموذج على النحو التالي:

$$Y_t = Y_{t-1} + e_t,$$

أو

$$\Delta Y = e_t.$$

لاحظنا بعد ذلك أن هناك أسباباً جيدة للاعتقاد بأن مثل هذا النموذج قد يكون ملائماً لقياس الظواهر الاقتصادية مثل الصدمات السعرية أو أسعار الصرف[1]. فالأحداث الفجائية تؤثر باستمرار على الصدمات السعرية (معكوسة في حد الخطأ e_t). وهذا يعني ضمناً أن المستثمرين لا يستطيعون توقع التغيرات المستقبلية في أسعار الأسهم. وإذا كان بإمكان المستثمرين فعل ذلك، فسوف تكون لديهم فرص لتحقيق أرباح. هناك القليل جداً من مثل هذه الفرص الربحية التي توجد في عالم الواقع (وإذا وجدت، فإنها تزول بسرعة لاستغلالها بواسطة المستثمرين البارعين).

يعد نموذج المسار العشوائي البسيط غير مناسب نوعاً ما كوصف للسلوك السعري للأسهم، إذ إن العديد من الأسهم ترتفع أسعارها مع مرور الزمن[2]. إن النموذج الأكثر واقعية إلى حد ما سيكون على النحو التالي:

$$\Delta Y_t = \alpha + e_t$$

فهذا النموذج يمكن تفسيره على أنه يعني ضمناً أن متوسط أسعار الأسهم يزداد بمعدل α لكل فترة ولكنها غير قابلة للتوقع. وهذا النموذج المعروف باسم

نموذج "المسار العشوائي مع اتجاه" (random walk with drift) يضيف قاطعاً لنموذج المسار العشوائي مما يتيح لأسعار الأسهم الاتجاه إلى أعلى مع مرور الزمن.

فإذا كانت فرضية المسار العشوائي صحيحة فسوف يتضح أن الاقتصاديين الماليين لا يستطيعون فعلاً دراسة أسعار الأسهم بطريقة عملية. إن الهدف من معظم الأبحاث التطبيقية بصفة عامة هو استخدام المتغيرات التفسيرية لتوضيح التباين في متغير تابع. ففي الحالة الراهنة، لا يمكن تفسير سلوك أسعار الأسهم عملياً سوى القول بأن تغيراتها غير قابلة للتوقع[3]. إذن ماذا يفعل الاقتصاديون الماليون المهتمون بسلوك أسعار الأسهم؟ تتمثل إحدى الإجابات في أنهم يحاولون تفسير تقلبات أسعار الأسهم، وأنهم مهتمون بصفة خاصة في معرفة ما إذا كان بالإمكان توقع تلك التقلبات (volatility).

وقد تكون مندهشاً حول أسباب اهتمامنا بتذبذبات المتغير. فالتقلبات تلعب دوراً مهماً في العديد من التطبيقات الاقتصادية الكلية والتمويل. لقد اهتمت الأعمال المبكرة في مجال الاقتصاد الكلي بتقلبات التضخم عبر الزمن، وانصب الاهتمام على معرفة ما إذا كانت تقلبات معدل التضخم قد تغيرت عبر الزمن (أي كانت هناك بعض الفترات المستقرة جداً التي ظل فيها التضخم ثابتاً تقريباً وكذلك فترات أخرى غير مستقرة). أوضحت بعض نظريات التوقعات الرشيدة (Rational Expectatiour) في الاقتصاد الكلي أن المهم في الاقتصاد ليس هو مستوى التضخم بل تباينه. حتى إذا كان مستوى التضخم مرتفعاً، فبالإمكان التخطيط للمستقبل بدرجة عالية من الثقة إذا كان التباين في التضخم منخفضاً. ورغم ذلك، فإن التباين المرتفع يعني صعوبة التوقع بمعدل التضخم خلال الفترة القادمة بدرجة دقيقة، وهذا له انعكاسات على العديد من نماذج الاقتصاد الكلي.

كمثال آخر لأهمية التقلبات في الاقتصاد، يمكننا النظر إلى سعر صرف الجنيه مقابل الدولار. فهذا السعر في غاية الأهمية من منظور الصادرات من المملكة المتحدة إلى الولايات المتحدة (والعكس). إن الشركات التي تنتج بغرض التصدير غالباً ما تحتاج لوضع خطط إنتاج تستند إلى توقعاتها لأسعار الصرف المستقبلية. فإذا كان سعر الصرف يعكس مستوى تقلبات منخفضة، سيصبح من السهل وضع الخطط في حين يكون ذلك صعباً في ظل درجة التقلب العالية. إن التأثير السلبي لعدم التأكد يسهم جزئياً في الجهود المتزايدة بين الدول لتثبيت أسعار الصرف أو تطبيق عملة مشتركة (مثل الاتحاد النقدي الأوروبي). كما يسهم أيضاً في العدد المتزايد من المشتقات المالية (مثل العقود المستقبلية والخيارات) التي يمكن استخدامها بواسطة الشركات لتجنب المخاطر الناجمة عن تقلبات أسعار صرف العملة.

ترتبط التقلبات في أسواق الأسهم أيضاً بدرجة المخاطرة. فإذا كان السهم يتصف بتذبذب عال فإن سعره قد يرتفع أو ينخفض بصورة كبيرة. بإمكان المستثمر الذي يرغب في شراء مثل هذا السهم المتقلب تحقيق أرباح كبيرة إذا ارتفع السعر بصورة كبيرة ولكنه أيضاً قد يخسر أمواله في حالة انخفاض السعر. وهذا الوضع يشير إلى أن درجة التقلبات تعد مقياساً للمخاطرة. ورغم ذلك، ينبغي علينا توخي الحيطة والحذر عند مضاهاة درجة التقلب بالمخاطرة. تؤكد النماذج المالية مثل نموذج تسعير الأصول الرأسمالية Capital Asset Pricing Model (CAPM) أن درجة المخاطرة لمحفظة الأسهم لا تعتمد على درجة تقلب الأسهم الفردية فحسب بل أيضاً على الارتباط بين أسهم المحفظة. دعنا نأخذ مثالاً متطرفاً: افترض أنك قد اشتريت محفظة مكونة من سهمين كانا منهما على درجة عالية من التقلب ولكنهما أيضاً مرتبطين بدرجة سالبة عالية. وهذا الارتباط السالب يعني أنه عندما ينخفض سعر أحد السهمين فإن سعر الآخر

يرتفع. لذلك، ورغم أن كل سهم في ذاته ينطوي على مخاطر (نظراً لدرجة التقلب العالية) إلا أن المخاطر توازن بعضها وتكون المحفظة في مجملها آمنة. إن أقصى ما يمكن قوله هنا هو أنه في الوقت الذي تمثل فيه تقلب الأسهم جانباً مهماً في القرار الاستثماري إلا أن هناك جوانب أخرى مهمة أيضاً ولكنها لا تخضع للنقاش هنا. سنفترض في باقي هذا الجزء أن نموذج المسار العشوائي لأسعار الأصول هو النموذج الصحيح. لذلك، سنفترض إما أن سعر الأصل المالي يتبع مساراً عشوائياً بحتاً أو يتبع مساراً عشوائياً مع اتجاه، وإننا قد أخذنا اشتقاقات من المتوسطات. ولتجنب الخلط والارتباك، سنجعل Δy_t يمثل سلسلة انحرافات من المتوسطات (أي $\Delta y_t = \Delta Y_t - \Delta \bar{Y}$ حيث إن $(\Delta \bar{Y} = \sum \Delta Y_t / T$). تذكر أن أخذ الانحرافات من المتوسط يعني ضمناً عدم وجود قاطع في النموذج (انظر الملحق رقم 4-1). وهكذا، حتى إذا كان سعر الأصل متجهاً إلى أعلى عبر الزمن، فإنه يمكن تجاهل شرط الاتجاه وكتابة $\Delta y_t = e_t$.

لقد استخدمنا حتى الآن مصطلح "تقلبات" على نحو عام، ولكن آن الأوان لتقديم تعريف منهجي.

إننا نستخدم Δy_t^2 كتقدير للتقلب في الزمن t. ولكن ينبغي ملاحظة أن التقلبات العالية تصاحبها تغيرات كبيرة إيجابية أو سلبية. ولما كان مربع أي رقم موجب القيمة، فإن الارتفاع الكبير أو الانخفاض الكبير في سعر الأصل يعني أن Δy_t^2 كبير وموجب. وعلى العكس، فإن سعر الأصل في أوقات الثبات لا يشهد

تغيراً كبيراً ويكون Δy_t^2 صغيراً. لذلك، سيكون مقياسنا للتقلب صغيراً في أوقات الثبات وكبيراً في أوقات التقلبات.

ويمكن الحصول على خيار بديل لقياسنا لمدى التقلبات وذلك من خلال الرجوع إلى بعض مواد الفصل الثاني. فقد تم التأكيد هناك على أن التباين هو مقياس تقلب المتغير. وبصفة عامة، يعد من الممارسات الشائعة الموازنة بين الاثنين واستخدام التباين مقياساً للتقلب. ولكن استخدام التباين مقياساً للتقلب يثير مشكلات في السياق الحالي. فهنا تبرز نقطة أساسية هي أننا نريد السماح لتقلب الأصل التغير عبر الزمن. إن التقلب في الفترة t قد يختلف عنها في الفترة t-1 أو t+1،... الخ. لقد استخدمنا في الفصل الثاني كافة المشاهدات لتوفير تقدير واحد للتباين. وهنا يمكننا استخدام مشاهدة واحدة في الفترة t لتوفير تقدير للتباين في الفترة t. (بعبارة أخرى، من غير المجدي استخدام بيانات الفترة t+1 لتقدير التباين في الفترة t نظراً لأن التباين قد يكون مختلفاً في الفترتين).

فإذا:

1) لاحظت أنه يمكننا استخدام مشاهدة واحدة فقط لتقدير التباين.

2) لاحظت أننا قد افترضنا أن البيانات عبارة عن انحرافات من المتوسط، ومن ثم فإن متوسطها يساوي صفراً.

3) استخدمت معادلة التباين من الفصل الثاني ثم حصلت على Δy_t^2 كتقدير للتباين[4].

يمكنك حساب هذا المقياس الخاص بتقلب سعر الأصل بسهولة في أي جدول إلكتروني أو برنامج حاسب آلي خاص بالاقتصاد القياسي من خلال إجراء فروقات بيانات سعر الأسهم بأخذ انحرافات من المتوسطات ثم تربيعها. بمجرد

عمل ذلك، ستحصل على متغير سلسلة زمنية جديد (وهو التقلب) يمكنك إجراء تحليل له باستخدام الأدوات المذكورة آنفاً.

يتم استخدام نماذج الانحدار الذاتي عادة لنمذجة "التركز (Clustering) في التقلب" الذي يحدث غالباً في بيانات السلاسل الزمنية المالية. اعتبر، على سبيل المثال، نموذج AR(1) يستخدم التقلب كمتغير سلسلة زمنية:

$$\Delta y_t^2 = \alpha + \phi \Delta y_{t-1}^2 + e_t$$

في هذا النموذج يعتمد التقلب في فترة على التقلب في فترة سابقة. فإذا كان $\phi > 0$ ، مثلاً، وإذا كانت درجة التقلب عالية بصورة غير عادية في الفترة السابقة (مثلاً، Δy_{t-1}^2 كان كبيراً جداً) فإنها ستكون عالية أيضاً بصورة غير عادية خلال هذه الفترة. وفي المقابل إذا كانت درجة التقلب منخفضة بصورة غير عادية (مثلاً، Δy_{t-1}^2 كان مساوياً للصفر تقريباً) فإن تقلب الفترة سوف تكون منخفضة أيضاً. بعبارة أخرى، إذا كان التقلب منخفضاً فسوف يميل إلى أن يظل منخفضاً، وإذا كان مرتفعاً سيميل إلى أن يظل مرتفعاً، وبالطبع فإن وجود الخطأ e_t يعنى أنه قد تكون هناك توقعات لهذا النمط. ولكن وبصفة عامة، يعنى هذا النموذج أننا سوف نميل إلى مشاهدة فترات أو تركزات عبر الزمن الذي تكون فيه درجة التقلب منخفضة وفترات حيث تكون فيها درجة التقلب مرتفعة. ويعد مثل هذا النمط شائعاً في الدراسات التطبيقية لأسعار الأصول. وكمثال في هذا الصدد، ارجع بالذاكرة إلى الفصل الثاني، الشكل رقم (2-1) (الرسم البياني لسعر صرف الجنيه مقابل الدولار). فإذا تأملت هذا الشكل مرة أخرى، فسوف تجد دورات طويلة تغير فيها سعر الصرف بصورة طفيفة (مثلا 1949-1967،

1993-1996) ودورات طويلة أخرى (1985-1992) حيث كان السعر أكثر تقلباً.

تختص المناقشة السابقة بنموذج AR(1) ولكن يمكن تعميمها على نموذج $AR(p)$. فكل النقاش الذي ورد في الفصل التاسع يعد وثيق الصلة بذلك. ولكن الفرق الوحيد هو أن التفسير يختص بتقلب السلسلة وليس بالسلسلة ذاتها. علاوة على ذلك، تعد كافة الأساليب الإحصائية التي ذكرناها في الفصل التاسع، وثيقة الصلة هنا. وباعتبار السلسلة مستقرة (مثلا، $|\phi|<1$ في حالة (AR(1))، يمكن تفسير تقديرات المربعات الصغرى الاعتيادية وقيم الاحتمال p بالطريقة المعتادة. كما يمكن إجراء اختبار التحقق من جذر الوحدة باستخدام اختبار ديكي - فولير. بإيجاز، ليس هناك ماهو جديد إحصائياً في هذا المقام.

مثال: تقلبات أسعار الأسهم:

يتضمن الملف STOCK.XLS بيانات عن Y= سعر سهم شركة تم جمعها أسبوعياً لفترة 4 سنوات (أي أن T=208). تم تحويل البيانات إلى لوغاريثمات. يوضح الشكل رقم (11-1) الرسم البياني للسلسلة الزمنية لهذه البيانات.

نلاحظ أن سعر هذا السهم قد مال للتزايد مع الزمن على الرغم من وجود فترات انخفض فيها سعر السهم. بلغ سعر السهم 24.53 جنيهاً في الأسبوع الأول وارتفع إلى 30.14 جنيهاً في الأسبوع 208 (بعبارة أخرى، $\ln(24.53)$ يساوي 3.200 و $\ln(30.14)$ يساوي 3.406).

يوضح الشكل رقم (11-2) الرسم البياني للنسبة المئوية للتغير في Y، أى ΔY. وحيث إن $100\times[\ln(Y_t)-\ln(Y_{t-1})]$ يمثل التغير النسبي في سعر

السهم، فإننا نضرب الفرق الأول في البيانات المستخدمة للحصول على الشكل رقم (11-1)، في العدد 100.

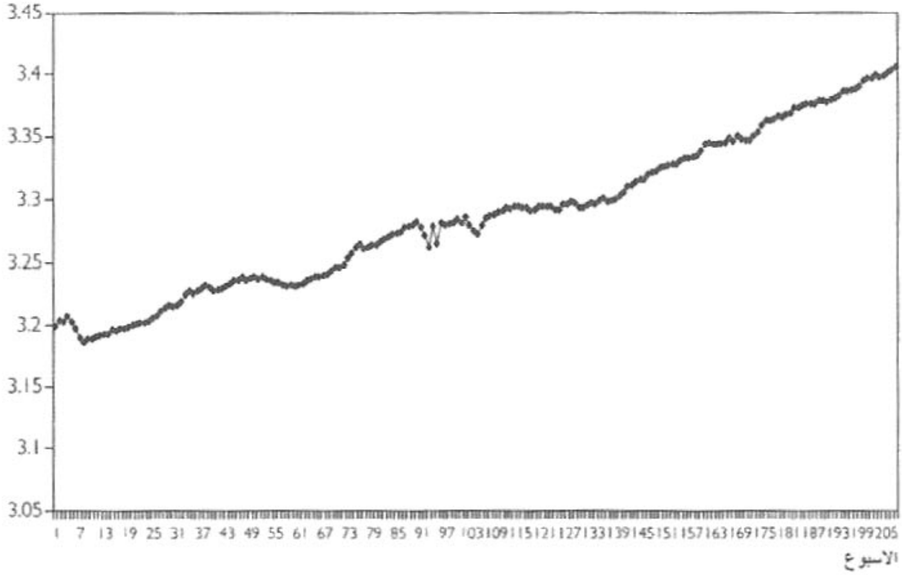
يشير هذا الشكل إلى أن التغير في سعر السهم في أي أسبوع كان موجباً ولكن السعر انخفض في بعض الأسابيع. وفي منتصف فترة الدراسة (الأسابيع 90-110)، حدثت تغيرات كبيرة عديدة (في الاتجاهين الموجب والسالب). ففي الأسبوعين 94 و 96 على سبيل المثال، ارتفع سعر السهم بأكثر من 1.5%. وهذه الزيادة تعد كبيرة في أسبوع واحد. فإذا كانت مثل هذه الزيادة قد استمرت لمدة عام، لكان سعر السهم قد تضاعف (أي أن العائد الأسبوعي البالغ 1.5% يتحول إلى عائد سنوي يزيد على 100%) ورغم ذلك، انخفض سعر السهم في الأسابيع 92-93-95 بنحو أكثر من ذلك. وبصفة عامة، كان سعر السهم في منتصف الفترة أكثر تقلباً منه في الفترات الأخرى.

لاختبار خصائص تقلب سعر السهم على نحو أكثر دقة، سنأخذ انحرافات من متوسط المشاهدات لبيانات الفروقات المستخدمة لإنشاء الشكل رقم (11-2) ومن ثم تربيعها. وذلك يعني أننا سوف:

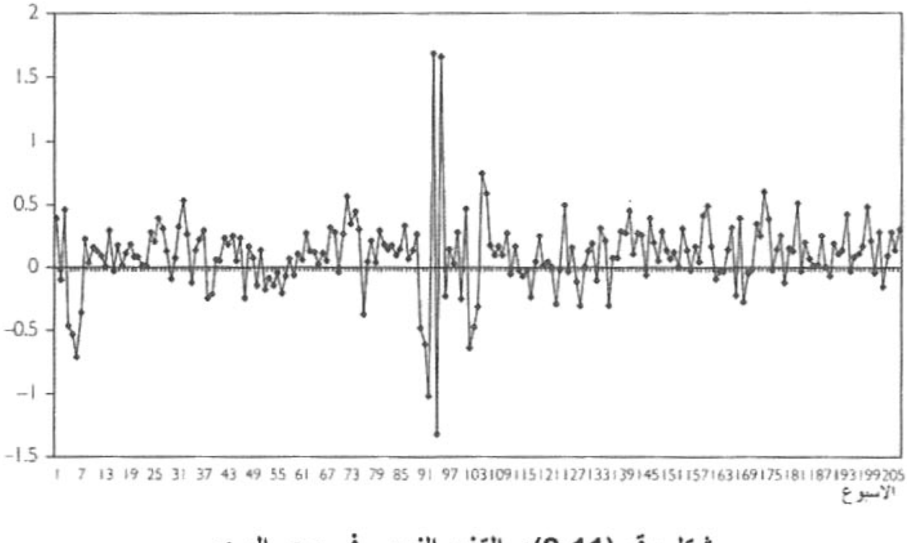
1) نحسب متوسط التغير في سعر السهم، 0.099%.

2) نطرح هذا الرقم من كل تغير في سعر السهم.

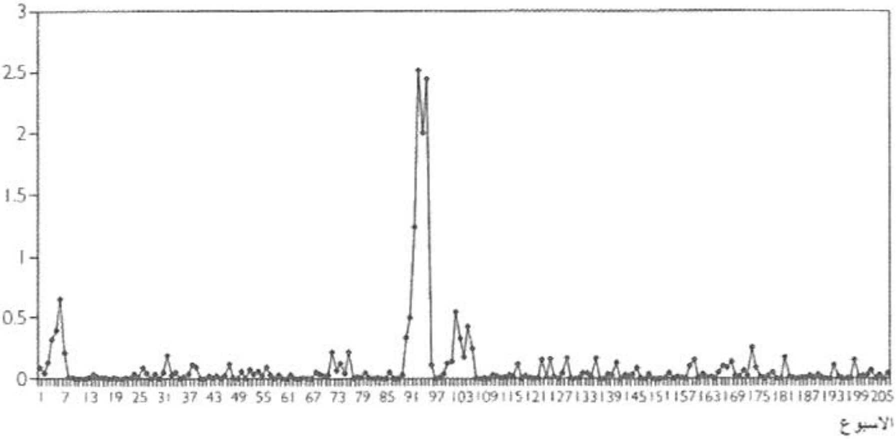
3) تربيع الناتج. يوضح الشكل رقم (11-3) عرضاً بيانياً للسلسلة الناتجة التي تمثل مقياس التقلب.



شكل رقم (11-1): لوغاريثم سعر السهم



شكل رقم (11-2): التغير النسبي في سعر السهم



شكل رقم (11-3): درجة تقلب سعر السهم

لاحظ أن التقلب هو مربع (Square) سعر السهم ومن ثم لا يمكن أن يكون سالباً. إن النمط الأكثر وضوحاً في الشكل رقم (11-3) هو الزيادة الكبيرة في مدى التقلبات في الأسابيع 90-97 وبدرجة أقل في الأسابيع 4-8 و 101-107. وهذا يوضح لنا بجلاء أن تقلب سعر هذا السهم يتزايد عبر الزمن.

ثمة دليل آخر أكثر أهمية لنمط التقلب يمكن الحصول عليه بإنشاء نموذج AR(p) باستخدام أساليب الفصل التاسع والتقلب كمتغير رئيس. يحقق إجراء الاختبار التتابعي المقترح في ذلك الفصل نموذج AR(1) الموضح في الجدول رقم (11-1). ويمكن مشاهدة أن التقلب في الأسبوع السابق له قوة تفسيرية قوية لتقلب هذا الأسبوع نظراً لأن معامله ذو معنوية إحصائية عالية. بالإضافة إلى ذلك، نلاحظ أن $R^2 = 0.54$ مما يدل على أن 54% من التباين في التقلب يمكن تفسيره بالتقلب في الفترة السابقة. ونتيجة لذلك لا يبدو أن هناك تركزاً عبر الزمن لهذه التقلبات. وهذه المعلومة قد تكون ذات أهمية بالغة للمستثمر الراغب في شراء هذا السهم.

جدول رقم (11-1): نموذج AR(1) باستخدام التقلب كمتغير رئيس

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
قاطع	0.024	0.015	1.624	0.106	-0.005	0.053
Δy_{t-1}^{2}	0.737	0.047	15.552	1.74E-36	0.643	0.830

افترض أن مستثمراً قد لاحظ أن $\Delta y_{t-1} = 0$ وبالتالي فإن $\Delta y_{t-1}^{2} = 0$. بعبارة أخرى، تغير سعر السهم بمقدار المتوسط في الفترة t-1. فالمستثمر مهتم بالتوقع بمدى التقلب في الفترة t من أجل تقدير المخاطر المحتملة في شراء السهم. ولما كان من غير الممكن توقع الخطأ، فإن المستثمر سوف يتجاهله (أي من المحتمل أن يكون موجباً أو سالباً). وفيما يلي نموذج AR(1) المقدر:

$$\Delta \hat{y}_{t}^{2} = 0.024 + 0.737_{t}\Delta y_{t-1}^{2}$$

وحيث إن $\Delta y_{t-1}^{2} = 0$، فإن المستثمر يتوقع بأن التقلب في الفترة t هي 0.024. ولكن إذا كان قد لاحظ أن $\Delta y_{t-1}^{2} = 1$، لكان قد توقع بدرجة تقلب في الفترة t مقدارها 0.761 (أي0.737+0.024). بالإمكان دمج مثل هذه المعلومات في النماذج المالية لسلوك المستثمر.

تمرين رقم (11-1):

يحتوي الملف NYSE.XLS بيانات عن ΔY = التغير النسبي في أسعار الأسهم في كل شهر للفترة 1952-1995 في بورصة نيويورك (NYSE). وللمهتمين في التفاصيل الدقيقة، فإن البيانات عبارة عن عوائد أسهم مرجحة حسب القيمة وخالية من عوائد الأسهم المخفضة باستخدام الرقم القياسي لأسعار السلع الاستهلاكية. لاحظ أن هذه البيانات في شكل فروقات ولكن لم يتم أخذ انحرافات من المتوسط، أي أنها ΔY وليس Y أو Δy.

أ) وضح بيانياً السلسلة الزمنية لهذه البيانات. دون ملاحظاتك على أي أنماط تشاهدها.

ب) باستخدام الأساليب الموضحة في الفصل التاسع، دون ملاحظاتك على خواص السلسلة الزمنية الأحادية لـ: ΔY. ما شكل دالة ارتباطه الذاتي؟ إذا أنشأت نموذج AR(p) باستخدام هذه البيانات، فكم تكون قيمة p؟ هل ΔY مستقرة؟ هل يمكن توقع عائدات الأسهم في بورصة نيويورك (أي، هل تساعدك عائدات الأسهم السابقة في توقع القيم الحالية)؟

ت) افترض أن السلسلة الأصلية Y، تتبع مسارات عشوائية بحيث يصبح نموذج AR(0) لـ ΔY ملائماً (ربما مع قاطع). احسب تقلب هذا المتغير حسب الموضح في هذا الفصل.

ث) وضح بيانياً تقلب هذه السلسلة. هل يوجد تركز في التقلبات؟

ج) أنشئ نموذج AR(p) لسلسلة التقلبات وناقش خواصه. هل بإمكان قيم التقلبات السابقة في سوق الأسهم مساعدتك في توقع درجة التقلب الحالية؟

سببية جرانجر (Granger Causality):

لقد أشرنا إلى سببية جرانجر في هذا الكتاب سريعاً وغالباً عبر التحذيرات بشأن تفسير نتائج الارتباط والانحدار على أنها تعكس علاقة سببية. لقد تناولنا في الفصل الثالث، على سبيل المثال، مثالاً يوضح أن معدلات تناول المشروبات الكحولية وسرطان الرئة مرتبطان، على الرغم من أن تناول مشروبات كحولية لا يسبب سرطان الرئة، إن الارتباط هنا لا يعني السببية. وفي الحقيقة، إن تدخين السجائر هو الذي يسبب سرطان الرئة، ولكن الارتباط بين تدخين السجائر وتناول المشروبات الكحولية أعطى علاقة واضحة بين تناول المشروبات الكحولية وسرطان الرئة.

عند مناقشتنا للانحدار، كنا نقف على أرضية أقل ثباتاً نظراً لمحاولتنا استخدام التبرير الاقتصادي والمنطق في تسمية متغير ما بالمتغير التابع وأخرى بالمتغيرات التفسيرية. وفي العديد من الحالات، ونظراً لأن الأخير يفسر الأول، كان من الملائم التحدث عن X "يسبب" Y (X Cause Y). فعلى سبيل المثال، أوضحنا في مثال أسعار المنازل في الفصول 4-7 أن سعر المنزل تتسبب فيه الخصائص المميزة للمنزل (مثلاً: عدد غرف النوم، عدد الحمامات،...الخ). ولكن، وعند مناقشتنا لتحيز المتغيرات المستبعدة في الفصل السادس، أصبح من الواضح أن الانحدار المتعدد قد يعطي تفسيراً مضللاً لدرجة السببية الموجودة في حالة استبعاد متغيرات تفسيرية مهمة. علاوة على ذلك، هناك العديد من نماذج الانحدار التي لايتضح فيها أي متغير يسبب الآخر. فعلى سبيل المثال، قمنا في الفصل العاشر (تمرين رقم 10-8) بإجراء انحدار Y= تضخم الأجور و X= تضخم الأسعار. من المحتمل أن يكون تضخم الأسعار سبباً لحدوث تضخم الأجور (أي أن X يسبب Y) نظرا لمطالبة العمال بأجور أعلى إذا كانت الأسعار متزايدة بسرعة. ورغم ذلك يمكن القول بأن Y تسبب X لأن زيادة الأجور تقلل

من أرباح الشركة ما لم تتم زيادة الأسعار. وهكذا يمكن لتضخم الأجور أن يتسبب في تضخم الأسعار. بعبارة أخرى، قد تحدث السببية في اتجاه واحد أو في اتجاهين. لذلك، فعند استخدام كلمة "سبب" مع نتائج الانحدار أو الارتباط يجب توخي الحيطة والحذر واستخدام المنطق.

ورغم ذلك، يمكننا مع بيانات السلاسل الزمنية، عمل عبارات أكثر قوة عن السببية من خلال الاستفادة من حقيقة أن الزمن لا يعود إلى الوراء. فإذا وقع الحدث A قبل الحدث B يمكن القول بأن A قد تتسبب في B. ولكن من المستحيل أن يكون B سبباً في حدوث A. وبعبارة أخرى، من المحتمل أن تكون أحداث الماضي سبباً في حدوث الأحداث الحالية، ولكن من المستحيل أن تفعل الأحداث المستقبلية ذلك.

بالإمكان فحص هذه الأفكار البديهية باستخدام نماذج الانحدار التي تتضمن سببية جرانجر. تتمثل الفكرة الأساسية في أن المتغير X يسبب Y حسب سببية جرانجر إذا كانت قيم X السابقة قادرة على المساعدة في تفسير Y. وبالطبع إذا ثبت صحة سببية جرانجر فإن ذلك لايضمن أن X قد سبب Y. وهذا هو سبب قولنا "سببية جرانجر" بدلاً من مجرد "سببية". ورغم ذلك، إذا كانت لقيم X السابقة قدرة تفسيرية لقيم Y الحالية، فإن ذلك يعني على الأقل أن X قد يكون سبباً في حدوث Y.

تعد سببية جرانجر ملائمة فقط لمتغيرات السلاسل الزمنية. ولتوضيح المفاهيم الأساسية، سننظر في سببية جرانجر بين متغيرين X و Y مستقرين. وسوف نتناول أدناه بإيجاز حالة غير مستقرة تتضمن X و Y ذات جذور وحدة ولكنهما متكاملان.

سببية جرانجر في نموذج ADL بسيط:

نظراً لافتراضنا أن X و Y مستقرة، فإن المناقشة الواردة في الفصل العاشر تشير إلى أن نموذج ADL ملائم. افترض أن نموذج ADL البسيط التالي يتم تطبيقه:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \beta_1 X_{t-1} + e_t.$$

إن هذا النموذج يعني أن قيمة X للفترة السابقة لها قدرة تفسيرية لقيمة Y الحالية. ويقيس المعامل β_1 تأثير X_{t-1} على Y_t. فإذا كان $\beta_1 = 0$، فلن تكون لقيم X السابقة أي تأثير على Y ولن يتسبب X في Y حسب سببية جرانجر. بعبارة أخرى، إذا كان $\beta_1 = 0$ فلن يتسبب X في Y حسب سببية جرانجر. وهناك بديل آخر للتعبير عن هذا المفهوم وهو "إذا كان $\beta_1 = 0$ فإن قيم X السابقة لن تكون ذات قدرة تفسيرية لـ Y أكثر من تلك التي تقدمها القيم السابقة لـ Y". وبما أننا نعرف كيف نقدر ADL ونجرى اختبارات الفروض، فمن السهل إجراء اختبار سببية جرانجر. أي أن تقدير المربعات الصغرى الاعتيادية للانحدار أعلاه يمكن الحصول عليه باستخدام أي برنامج حاسب آلي خاص بالاقتصاد القياسي واختبار معنوية قيمة الاحتمال p للمعامل الخاص بالمتغير X_{t-1}. فإذا كان المعامل β_1 ذا معنوية إحصائية (مثلاً، قيمة الاحتمال p أقل من 0.05) سنخلص إلى أن X يسبب Y حسب سببية جرانجر. لاحظ أن فرض العدم الذي خضع للاختبار هو $H_0: \beta_1 = 0$ (لن تحدث سببية جرانجر). لذلك يجب الإشارة إلى اختبار $\beta_1 = 0$ بأنه اختبار "عدم سببية جرانجر" (Granger non causality) ولكننا سنشير إلى هذا الإجراء بـ"اختبار سببية جرانجر" (Granger causality test) وهو الأكثر شيوعاً.

سببية جرانجر في نموذج ADL بفترات تباطؤ: q,p:

يعد نموذج ADL أعلاه محدوداً لتضمنه متباطئة واحدة لـ: X وY. سنحتاج، بصفة عامة، إلى اختيار قيم المتباطئة باستخدام الأساليب الموضحة في الفصل العاشر ADL(p,q) بالشكل التالي[5]:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + \beta_1 X_{t-1} + ... + \beta_q X_{t-q} + e_t$$

هنا، X يسبب Y حسب سببية جرانجر إذا كان إحدى المعاملات $\beta_1,......,\beta_q$ أو جميعهم ذات معنوية إحصائية. بعبارة أخرى، إذا كان X في أي وقت في الماضي ذا قدرة تفسيرية لقيمة Y الحالية، فإننا نقول إن X يسبب Y حسب سببية جرانجر. ولأننا نفترض أن Y و X لا يحتويان على جذور وحدة، فبالإمكان استخدام تحليل انحدار المربعات الصغرى الاعتيادية في تقدير هذا النموذج. ويمكن استخدام قيم P للمعاملات المفردة للتحقق من وجود سببية جرانجر. فإذا كنت تستخدم مستوى معنوية مقداره 5%، و كانت أي من قيم P للمعاملات $\beta_1,...,\beta_q$ أقل من 0.05، فسوف تستنتج وجود سببية جرانجر. أما إذا لم تكن أي من قيم P أقل من 0.05، فسوف تستنتج عدم وجود سببية جرانجر.

تُعد الطريقة الموضحة أعلاه مفيدة ويمكن تنفيذها بسهولة ببرنامج إكسل أو أي برنامج حاسب آلي إحصائي. من المحتمل الحصول على دليل مؤكد حول ما إذا كان X يسبب Y حسب سببية جرانجر. ولكن يجب ملاحظة أن هناك طريقة أكثر صحة (ولكنها أكثر تعقيداً) لإجراء هذا الاختبار . تذكر أن فرض العدم الخاضع للاختبار هو عدم سببية جرانجر. أي أن X لا يسبب Y حسب سببية جرانجر إذا لم تكن لقيم X السابقة قدرة تفسيرية لقيمة Y الحالية. سنحتاج بعد ذلك لاختبار فرض العدم $H_0: \beta_1 = \beta_2 = ... = \beta_q = 0$. والوصول إلى أن X يسبب Y

حسب سببية جرانجر فقط في حالة رفض فرض العدم. لاحظ أن هذا الاختبار يختلف قليلاً عن ذلك المقترح في الفقرة السابقة. أي أن الاختبار المشترك لـ: $\beta_1 = \beta_2 = ... = \beta_q = 0$ لايساوي تماماً عدد q اختباراً مفرداً $i = 1,...,q$. لم نناقش كيفية إجراء الاختبارات لتحديد ما إذا كانت عدة معاملات مجتمعة تساوي صفراً. بالنسبة للقراء المهتمين بمثل هذه الاختبارات المشتركة، يمكن الرجوع إلى الملحق رقم (11-1) الذي يتضمن بعض الإرشادات العملية. ولكن إذا فضلت اتباع الطريقة البسيطة الموضحة أعلاه، فلا بد من ملاحظة ما يلي:

"إذا وجدت أن أياً من المعاملات $\beta_1,...,\beta_q$ ذات معنوية باستخدام إحصائية t أو قيم الاحتمال p للمعاملات المفردة، فبإمكانك الاستنتاج بسهولة أن X يسبب Y. أما إذا لم يكن أي من هذه المعاملات ذات معنوية، فمن المحتمل أن X يسبب Y حسب سببية جرانجر. ورغم ذلك، فإنك ستكون أكثر احتمالاً للخطأ إذا استنتجت الأخير أكثر من استخدامك الاختبار المشترك الصحيح لعدم سببية جرانجر.

مثال: هل يؤدى تضخم الأجور إلى تضخم الأسعار حسب سببية جرانجر؟

يحتوى الملف WP.XLS بيانات عن الأسعار والأجور في المملكة المتحدة خلال الفترة 1855-1987 (تمرين رقم 10-8). فإذا كنت قد قمت بحل ذلك التمرين، فسوف تتذكر أن لوغاريثمات الأجور والأسعار أظهرت جذور وحدة ولكنها لم تظهر وجود تكامل. ولكن فروقات هذه السلاسل كانت مستقرة ويمكن اعتبارها بمثابة معدلات تضخم (أي تضخم الأجور وتضخم الأسعار). سنستخدم هذه البيانات للتحقق عما إذا كان تضخم الأجور في الماضي يسبب تضخم أسعار. هناك سبب جيد للاعتقاد بأن هذا هو الوضع الفعلي.

عموماً، إذا كانت الأجور متزايدة، ستعمل الشركات على زيادة الأسعار للمحافظة على هوامش الربح عند مستوياتها الحالية.

يوضح الجدول رقم (2-11) نتائج من تقدير المربعات الصغرى الاعتيادية للانحدار ΔP =تضخم الأسعار مع أربع متباطئات ذاتية وأربع متباطئات لـ ΔW = تضخم الأجور واتجاه عام غير عشوائي.

جدول رقم (2-11): نموذج ADL باستخدام تضخم الأسعار كمتغير تابع

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
قاطع	-0.751	0.710	-1.058	0.292	-2.156	0.654
ΔP_{t-1}	0.822	0.170	4.850	3.81E-6	0.486	1.158
ΔP_{t-2}	-0.041	0.186	-0.222	0.825	-0.409	0.326
ΔP_{t-3}	0.142	0.186	0.762	0.448	-0.227	0.511
ΔP_{t-4}	-0.181	0.175	-1.035	0.303	-0.526	0.165
ΔW_{t-1}	-0.016	0.143	-0.114	0.903	-0.299	0.267
ΔW_{t-2}	-0.118	0.143	-0.823	0.412	-0.402	0.166
ΔW_{t-3}	-0.042	0.143	-0.292	0.771	-0.324	0.241
ΔW_{t-4}	0.038	0.142	0.266	0.791	-0.244	0.319
الوقت (Time)	0.030	0.011	2.669	0.009	0.0077	0.052

يشير فحص قيم الاحتمال p في الجدول إلى أن الاتجاه غير العشوائي وتضخم الأسعار للفترة السابقة فقط يتمتعان بقدرة تفسيرية للتضخم الحالي. وتعد كافة معاملات متباطئات تضخم الأجور غير معنوية. وبعكس توقعاتنا، لايبدو أن تضخم الأجور يحدث تضخم الأسعار حسب سببية جرانجر. (يعتمد هذا الاستنتاج على فحص قيم الاحتمال p المفردة لكل معامل. يوضح الملحق رقم (11-1) تفاصيل الاختبار المشترك لـ: $\beta_1 = \beta_2 = ... \beta_4 = 0$ ويدعم النتيجة المتمثلة في أن تضخم الأجور لايسبب تضخم الأسعار حسب سببية جرانجر).

السببية في الاتجاهين:

ليس من الواضح في العديد من الحالات، في أي اتجاه ينبغي أن تتم السببية. فعلى سبيل المثال، هل يجب أن يسبب تضخم الأجور السابق تضخماً سعرياً أم أن العكس يجب أن يتم؟ في مثل هذه الحالات التي قد تكون السببية في الاتجاهين، يجب عليك اختبار ذلك. وإذا كان Y و X هما المتغيران قيد الدراسة، فإنه بالإضافة إلى إجراء انحدار Y على متباطئاته الذاتية ومتباطئات X (كما أعلاه)، يجب عليك إجراء انحدار X على متباطئاته الذاتية ومتباطئات Y.

لاحظ أنه بالإمكان الوصول إلى أن Y يسبب X حسب سببية جرانجر وأن X أيضا يسبب Y. في حالة النماذج الاقتصادية المعقدة، تعد السببية في الاتجاهين شيئا مألوفاً وحتى معقولاً. قد تفكر، على سبيل المثال، في العلاقة بين أسعار الفائدة وأسعار الصرف. فمن المنطقي، من المنظور الاقتصادي الكلي، القول بأن سياسة سعر الفائدة قد تؤثر في أسعار الصرف المستقبلية. ومن ناحية أخرى، من المنطقي أيضا القول بأن أسعار الصرف قد تؤثر أيضاً في سياسة سعر الفائدة المستقبلية (مثلاً: إذا اعتبر سعر الصرف مرتفعاً جداً في الوقت الحالي، فإن البنك المركزى سيضطر إلى تخفيض أسعار الفائدة في المستقبل).

مثال: هل يؤدى تضخم الأسعار إلى تضخم الأجور حسب سببية جرانجر؟

لقد استخدمنا في المثال السابق ملف البيانات WP.XLS للتحقق عما إذا كان تضخم الأجور يسبب تضخم الأسعار حسب سببية جرانجر وتوصلنا إلى أن ذلك لايحدث. ولكن من الممكن أن تتم السببية في الاتجاه المعاكس. أي أن تضخم الأسعار قد يسبب تضخم الأجور. عموماً، يُرجع العمال والنقابات غالباً التضخم إلى الفترة السابقة وتتم المطالبة بزيادة الأجور تبعاً لذلك.

يحتوى الجدول رقم (11-3) نتائج تقدير المربعات الصغرى الاعتيادية للانحدار ΔW = تضخم الأجور على أربع متباطئات ذاتية وأربع متباطئات لتضخم الأسعار (ΔP) واتجاه عام غير عشوائي.

جدول رقم (11-3): نموذج ADL باستخدام تضخم الأجور كمتغير تابع

	معامل	الخطأ المعياري	إحصائية t	قيمة الاحتمال	الدنيا 95%	العليا 95%
قاطع	-0.609	0.835	-0.730	0.467	-2.262	1.044
ΔW_{t-1}	0.053	0.168	0.312	0.755	-0.280	0.386
ΔW_{t-2}	-0.040	0.169	-0.235	0.814	-0.374	0.294
ΔW_{t-3}	-0.058	0.168	-0.348	0.728	-0.391	0.274
ΔW_{t-4}	0.036	0.167	0.215	0.830	-0.295	0.367
ΔP_{t-1}	0.854	0.200	4.280	3.83E-5	0.459	1.249
ΔP_{t-2}	-0.217	0.218	-0.993	0.323	-0.649	0.215
ΔP_{t-3}	0.234	0.219	1.067	0.288	-0.200	0.668
ΔP_{t-4}	-0.272	0.205	-1.323	0.188	-0.678	0.135
الوقت(Time)	0.046	0.013	3.514	0.001	0.020	0.072

وهنا نجد دليلاً على أن تضخم الأسعار يسبب تضخم الأجور وفقاً لسببية جرانجر. وعلى وجه الخصوص، يعد المعامل الخاص بالمتغير ΔP_{t-1} ذا معنوية عالية مما يشير إلى أن معدل تضخم الأسعار للسنة السابقة له قدرة تفسيرية قوية لتضخم الأجور.

تمرين رقم (11-2):

باستخدام ملف البيانات WP.XLS في الأمثلة السابقة، قررنا أن $p = q = 4$ (أى أربع متباطئات لكل من تضخم الأجور وتضخم الأسعار).

باستخدام تضخم الأسعار كمتغير تابع وإجراء الاختبار الموضح في الفصل العاشر، اختر القيم المثلى لكل من p,q . وضح ما إذا كان تضخم الأجور يسبب تضخم الأسعار، باستخدام نموذج $ADL(p,q)$ الذي اخترته.

كرر التحليل باستخدام تضخم الأجور كمتغير تابع.

تمرين رقم (11-3):

يتضمن ملف إكسل LONGGDP.XLS بيانات سنوية عن نصيب الفرد من إجمالي الناتج المحلي الحقيقي لأربع من أكبر الدول الناطقة بالإنجليزية (الولايات المتحدة، المملكة المتحدة، كندا وأستراليا) للفترة 1870-1993.

أ) استخدم الفروقات للحصول على السلسلة الزمنية للنمو في نصيب الفرد من إجمالي الناتج المحلي لكل من الدول الأربعة.

ب) تحقق ما إذا كان نمو إجمالي الناتج المحلي في أى دولة يسبب النمو، في أى دولة أخرى حسب سببية جرانجر. مثلا، هل النمو في إجمالي الناتج المحلي في الولايات المتحدة يسبب النمو في إجمالي الناتج المحلي في المملكة المتحدة حسب سببية جرانجر؟ وهل يسبب ذلك في كندا؟

ركزت المناقشة المختصرة لسببية جرانجر على متغيرين: Y و X. ولكن ليس هناك أى سبب يحول دون توسيع هذه الأساليب الأساسية لتشمل حالة تتضمن العديد من المتغيرات. فإذا كانت لدينا، مثلاً، ثلاثة متغيرات: X و Y و Z وكنا نود التحقق عما إذا كان X و Z يسببان Y حسب سببية جرانجر، فسوف نقوم بسهولة بإجراء انحدار Y على متباطئات Y ومتباطئات X ومتباطئات Z. فإذا اتضح أن متباطئات Z، مثلاً كانت معنوية وأن متباطئات X ليست كذلك، يمكن القول أن Z يسبب Y حسب سببية جرانجر ولكن X لا يسبب Y.

سببية جرانجر مع وجود متغيرات تتصف بالتكامل المشترك:

إن التحقق من سببية جرانجر بين متغيرات تتميز بالتكامل المشترك يعد مماثلاً للطريقة الموضحة أعلاه. فمن المعتاد التعامل مع نموذج تصحيح الخطأ (ECM)، الموضح في الفصل العاشر:

$$\Delta y_t = \varphi + \delta t + \lambda e_{t-1} + \gamma_1 \Delta y_{t-1} + \ldots + \gamma_p \Delta Y_{t-p} + \omega_1 \Delta X_{t-1} + \ldots + \omega_q \Delta X_{t-q} + \varepsilon_t.$$

كما لاحظنا في الفصل العاشر، يعد هذا نموذج ADL باستثناء وجود الحد λe_{t-1}. تذكر أن $e_{t-1} = Y_{t-1} - \alpha - \beta X_{t-1}$ يمثل التقدير الذي يمكن الحصول عليه بإجراء انحدار Y على X وحفظ البواقي. يتسبب X في Y حسب سببية جرانجر إذا كانت قيم X السابقة ذات قدرة تفسيرية لقيم Y الحالية. وبتطبيق هذا

الفهم على نموذج تصحيح الخطأ، يمكننا ملاحظة أن قيم X السابقة تظهر في الحدود $\Delta X_{t-1},...,\Delta X_{t-q}$ بالإضافة إلى e_{t-1}. وهذا يعني ضمناً أن X لايسبب Y وفقاً لسببية جرانجر إذا كان $\omega_1 = ... = \omega_q = \lambda = 0$. يوضح الفصل العاشر كيف يمكننا استخدام انحداري المربعات الصغرى الاعتيادية لتقدير نماذج تصحيح الخطأ ومن ثم استخدام قيم الاحتمال p أو فترات الثقة للتحقق من وجود سببية. لذلك، يمكن استخدام إحصائية t وقيم الاحتمال p لاختبار سببية جرانجر بالطريقة ذاتها التي تمت مع حالة الاستقرار. ويمكن أيضاً استخدام F الموضح في الملحق رقم (11-1) لإجراء اختبار منهجي لفرض العدم:

$$H_0: \omega_1 = ... = \omega_q = \lambda = 0.$$

لقد ناقشنا في الفقرة السابقة كيفية إجراء اختبار للتحقق عما إذا كان X يسبب Y وفقاً لسببية جرانجر. ويمكن اختبار عما إذا كان Y يسبب X وفقاً لسببية جرانجر بعكس أدوار كل من X وY في نموذج تصحيح الخطأ. ومن النتائج المهمة لنظرية جرانجر التي تجدر الإشارة إليها مرة أخرى هنا (دون إثبات) أنه إذا كانت X وY تتميزان بالتكامل المشترك فإن شكلاً من أشكال سببية جرانجر سوف يحدث، وهذا يعني إما أن X يجب أن يسبب Y أو أن Y يجب أن يسبب X (أو الاثنين معاً).

تمرين رقم (11-4):

استخدم بيانات Y= الاستهلاك و X= الدخل الشخصي وذلك من الملف INCOME.XLS. افترض (ربما بصورة غير واقعية في ضوء تمرين رقم 10-5) أن كلاً من Y و X تتميزان بالتكامل المشترك. أجر اختباراً لتحديد ما إذا كانت Y تسبب X. أجر اختباراً عما إذا كانت X تسبب Y.

متجه الانحدار الذاتي:

من الطبيعي أن تؤدى مناقشتنا لسببية جرانجر إلى موضوع متجه الانحدارات الذاتية (VARs). وقبل مناقشة مدى شيوعها وتقديرها، سنقوم أولاً بتعريف متجه الانحدار الذاتي (VAR). سنفترض مبدئياً أن كافة المتغيرات مستقرة. فإذا كانت المتغيرات الأصلية ذات جذور وحدة، فسوف نفترض أن الفروقات تم أخذها بحيث يحتوي النموذج على التغيرات في المتغيرات الأصلية (التى لاتحتوي جذور وحدة). سيتناول آخر هذا الجزء توسع في هذه الحالة لتغطي حالة التكامل.

عند فحصنا لسببية جرانجر بين X وY، بدأنا بنسخة مقيدة من نموذج $ADL(p,q)$ مع Y كمتغير تابع. واستخدمنا ذلك في التحقق عما إذا كان X يسبب Y. ثم ذهبنا لتناول السببية في الاتجاهين الأمر الذي تطلب عكس دوري X وY في نموذج ADL، أى أصبح X المتغير التابع.

بإمكاننا كتابة المعادلتين كما يلي:

$$Y_t = \alpha_1 + \delta_1 t + \phi_{11} Y_{t-1} + \dots + \phi_{1p} Y_{t-p} + \beta_{11} X_{t-1} + \dots + \beta_{1q} X_{t-q} + e_{1t},$$
$$X_t = \alpha_2 + \delta_2 t + \phi_{21} Y_{t-1} + \dots + \phi_{2p} Y_{t-p} + \beta_{21} X_{t-1} + \dots + \beta_{2q} X_{t-q} + e_{2t}.$$

تختبر المعادلة الأولى ما إذا كان X يسبب Y، وتختبر الثانية ما إذا كان Y يسبب X حسب سببية جرانجر. لاحظ أن المعاملات الآن تحمل رموزاً تشير إلى أي من المعادلتين توجد فيها. فعلي سبيل المثال، يمثل α_1 القاطع في المعادلة الأولى في حين يمثل α_2 القاطع في المعادلة الثانية. علاوة على ذلك تحمل الأخطاء الآن رموزاً لتوضيح حقيقة أنها ستكون مختلفة في المعادلتين.

هنـاك معادلتـان تنطويـان علـى متجـه (Vector) انحـدار ذاتـي. ويعـد متجـه الانحدار الذاتي توسعاً لنموذج الانحدار إلى الحالة التى تتضمن أكثر من متغير. تذكر أن الانحدار الذاتي الموضح في الفصل التاسع يحتوي متغيراً واحداً تابعاً (Y_t) يعتمد على متباطئاته الذاتية (وربما اتجاه عام غير عشوائي). ولكن متجه الانحدار الذاتي للكميات المتجهة يتضمن أكثر من متغير تابع (Y وX مثلاً) ومن ثم هناك أكثر من معادلة واحدة (مثلاً، معادلة يكون فيها Y متغيراً تابعاً وأخرى يكون فيها X متغيراً تابعاً). تستخدم كل معادلة متباطئاتها الذاتية لكافة المتغيرات قيد الدراسة (وربما الاتجاه العام غير العشوائي).

تشكل المعادلتان أعلاه متجه انحدار ذاتي لمتغيرين. بإمكانك مثلاً أن تشاهد في المعادلة الأولى أن Y يعتمد على p متباطئة ذاتية و q متباطئة تخص X. ويمكن اختيار قيم المتباطئة p,q باستخدام أساليب الاختبار التتابعي الموضحة في الفصل التاسع. ورغم ذلك، وعلى وجه الخصوص إذا كان متجه الانحدار الذاتي يتضمن أكثر من متغيرين، فيجب اختيار قيم متباطئة مختلفة (أي واحد لكل متغير في المعادلة). وفي ضوء ذلك من الشائع تحديد $p=q$ واستخدام قيمة المتباطئة ذاتها لكل متغير في كل معادلة. ويطلق على النموذج الناتج نموذج متجه الانحدار الذاتي $VAR(p)$. يحتوي نموذج $VAR(p)$ التالي ثلاثة متغيرات: Y وX وZ:

$$Y_t=\alpha_1+\delta_1 t+\phi_{11}Y_{t-1}+\ldots+\phi_{1P}Y_{t-p}+\beta_{11}X_{t-1}+\ldots+\beta_{1P}X_{t-P}+\delta_{11}Z_{t-1}+\ldots+\delta_{1P}Z_{t-P}+e_{1t};$$

$$X_t=\alpha_2+\delta_2 t+\phi_{21}Y_{t-1}+\ldots+\phi_{2P}Y_{t-p}+\beta_{21}X_{t-1}+\ldots+\beta_{2P}X_{t-P}+\delta_{21}Z_{t-1}+\ldots+\delta_{2P}Z_{t-P}+e_{2t};$$

$$Z_t=\alpha_3+\delta_3 t+\phi_{31}Y_{t-1}+\ldots+\phi_{3P}Y_{t-p}+\beta_{31}X_{t-1}+\ldots+\beta_{3P}X_{t-P}+\delta_{31}Z_{t-1}+\ldots+\delta_{3P}Z_{t-P}+e_{3t}.$$

لاحظ أن كل معادلة تتضمن، بالإضافة إلى القاطع والاتجاه العام غير العشوائي، عدد p متباطئة لكل المتغيرات في الدراسة. بالإمكان الحصول على نماذج $VAR(p)$ مع أكثر من ثلاثة متغيرات بطريقة مماثلة.

نظراً لافتراضنا أن كافة المتغيرات في $VAR(p)$ مستقرة، بالإمكان إجراء التقدير والاختبار بطريقة قياسية، أي يمكنك الحصول على تقديرات المعادلات في كل معادلة باستخدام طريقة المربعات الصغرى الاعتيادية. بعد ذلك ستتيح قيم الاحتمال p أو إحصائيات t إمكانية التحقق عما إذا كانت المعاملات معنوية. بإمكانك أيضا استخدام الشرح الوارد في الملحق (11-1) لإجراء اختبارات F الأكثر تعقيداً.

تعد نماذج متجه الانحدارات الذاتية سهلة الاستخدام. ومع ذلك قد تكون مندهشاً حول ما الذي يدعونا للتعامل مع مثل هذه النماذج. يتمثل أحد الأسباب في اختبار سببية جرانجر، وهذا يعني أن نماذج VAR توفر إطاراً لاختبار سببية جرانجر بين كل مجموعة من المتغيرات. ورغم ذلك هناك أسباب أعمق لرغبتنا في استخدام هذه النماذج والتي ينبغي علينا ذكرها.

لقد ركزنا في هذا الكتاب كله على الحاجة إلى الحرص عند تفسيرنا نتائج الارتباط أو الانحدار، على أنها تعكس السببية أو التأثير. بالإمكان الاستفادة من النظرية الاقتصادية والمنطق في العديد من الحالات. لقد تناولنا في الفصلين الرابع والسادس العديد من الأمثلة والتي يمكن القول بشأنها أن الانحدارات تعكس السببية. فعلى سبيل المثال، تسبب X (الكثافة السكانية) في Y (انحسار الغابات) أو أن للمتغير X (مساحة قطعة الأرض) أثر في Y (سعر المنزل). ومن غير المعقول في كلتا الحالتين القول بأن Y أثر في X أو تسبب فيه.

ولكـن هنـاك العديـد مـن الحـالات التـي لاتجـدي فيهـا النظريـة الاقتـصادية أو المنطق في توفير نموذج انحدار يمكن تفسيره على أنه يعكس السببية. فعلى سبيل المثـال، هـل يـسبب Y (تـضخم الأجـور) X (تـضخم الأسـعار)؟ أو هـل يحـدث العكس؟ فالنظريـة الاقتـصادية والبديهـة تـدلان علـى أن كليهمـا قـد يحـدث وأن اختبارات سببية جرانجر قادرة على إلقاء الضوء على هذه الأسئلة. إن مجال الاقتصاد الكلـي خاصـة مزدحم بمثل هذه الأمثلـة. هل تتسبب أسعار الفائدة فـي تغير أسعار الصرف أو العكس أو كلاهما؟ هل يتسبب نمو إجمالي الناتج المحلي في تغير أسعار الفائدة؟ هل يصح العكس أو كلاهما؟ إن الإجابات غير واضحة ومن الصعب معرفة كيفية تفسير المعاملات في انحدار Y_t على X_t.

لقد تجاهلنا حتى الآن قضايا التكامل والمضاعف بعيد المدى. ولكن وحتى إذا كـان التكامل موجوداً، يجب أن نكون على قدر من الحرص عند تفسير نتـائج الانحدار على أنها تعكس السببية. لقد وجدنا في الفصل العاشر، مثلاً، أن أسعار Y (البرتقـال العـضوي) وأسـعار X (البرتقـال العـادي) متكاملـة ذاتيـاً وأن تـأثير المضاعف بعيد المدى لـ X على Y هو 0.966. ومن المحتمل أن تشير هذه النتائج إلى أن X قد أثر في Y (بمعنى أنـه إذا ارتفع سعر البرتقال العادي بمقدار بنس واحد، فمن المحتمل أن يرتفع سعر البرتقال العضوي بمقدار 0.966 بنساً على المدى البعيد). ولكن من غير المحتمل أن يؤثر سعر البرتقال العضوي على سعر البرتقال العادي لأن النوع الأول لايشكل سوى شريحة صغيرة من السوق. لذلك، فـإن X يـؤثر فـي Y ولكـن Y لايـؤثر فـي X. فـإذا كنـا قـد قلبنـا الأوضـاع وأصبح سعر البرتقال العادي هو المتغير التابع وسعر البرتقال العضوي هو المتغير التفسيري، لكنا قد وجدنا تكاملاً ذاتياً وحسبنا المضاعف بعيد المدى، ولكننا كنا سنقع في الخطأ إذا استخدمناه كمقياس لتأثير سعر البرتقال العضوي على سعر البرتقال العادي.

إن القضايا المثارة في الفقرات السابقة إما أنها لا تحدث إطلاقاً أو تحدث على نحو أقل بكثير بالمقارنة مع نماذج VAR. أي أن كافة المتغيرات التى استخدمناها لتفسير القيمة الحالية للمتغير التابع قد حدثت في الماضي (مثلاً، كانت كافة المتغيرات التفسيرية في المعادلة الأولى بتاريخ t-1 أو قبله بينما كان المتغير التابع Y_t). من المحتمل أن يؤثر الماضي في الحاضر ولكن من المستحيل أن يؤثر الحاضر في الماضي. لذلك، بإمكان المتغير التفسيري في نموذج VAR التأثير في المتغير التابع ولكن من غير الممكن أن يؤثر المتغير التابع في المتغير التفسيري. إن مشكلات التفسير التى تبرز مع انحدار Y_t على X_t لا تحدث في حالة نموذج VAR [6].

من الجوانب الجدلية المصاحبة لنماذج VAR أنها غير مبنية على نظريات (atheoretical)، أى أنها لاتعتمد كثيراً على النظرية الاقتصادية. فالنظرية تقتصر على اختبار المتغيرات في نموذج VAR اعتبر علي سبيل المثال، العلاقة بين أسعار الفائدة ومستوى الأسعار والعرض من النقود وإجمالي الناتج المحلي الحقيقي. لقد ناقش أصحاب نظريات الاقتصاد الكلي العديد من النماذج المتطورة لهذه العلاقة. إن نموذج IS-LM، الذي تم توسعته ليشمل التضخم، هو الأفضل تقريباً ولكن هناك نماذج متعددة غير نموذج IS-LM2. في الوقت الذي يرغب فيه المنظر الاقتصادي الكلي من هذه النظرية أن يكون لها أثر واضح على أرض الواقع، فإن ممارس نموذج VAR لايعول عليها إطلاقاً. ينص نموذج VAR على أن "أسعار الفائدة، مستوى الأسعار، العرض من النقود وإجمالي الناتج المحلي الحقيقي مترابطة. إننا نحاول نمذجة هذه العلاقة على أنها تعني ضمناً فقط أن كل متغير يعتمد على متباطئاته الذاتية وكافة المتغيرات الأخرى". ليس هناك أى علاقة حقيقية بين VAR التطبيقي ونموذج الاقتصاد الكلي النظري (نموذج IS-LM مثلاً).

سيدافع مستخدم نموذج VAR عنه من خلال ذكر أدائه الممتاز في مجال إجراء التوقعات المستقبلية. سنناقش هذه الميزة بتفصيل أكثر أدناه ولكن ينبغي عليك الآن معرفة أن ذلك يمثل سبباً رئيساً لاستخدام هذه النماذج. لقد اتضح في العديد من الحالات أن نماذج VAR تتمتع بقدرة أفضل لإجراء التوقعات بالمقارنة مع نماذج الاقتصاد الكلي المتطورة الأخرى. إن حقيقة أن طرق الانحدار البسيط التى تستخدم برامج الحاسب الآلي تتفوق في الأداء على النماذج الاقتصادية الكلية المعقدة المصممة والمدعومة من قبل المتخصصين في القطاع العام أو الخاص، تمثل دافعاً قوياً لاستخدام نماذج VAR.

مثال: نموذج متجه انحدار ذاتي $VAR(1)$ مع متغيرات RMPY:

يستخدم الاقتصاديون غالباً المتغيرات الاقتصادية الكلية المهمة التالية:

R= سعر الفائدة

M= عرض النقود

P= مستوى الأسعار

Y= إجمالي الناتج المحلي الحقيقي.

ونتيجة لهذه الرموز، يشار إلى النماذج التى تستخدم هذه المتغيرات بنماذج رمبي (RMPY).

يتضمن الملف RMPY.XLS بيانات ربع سنوية عن متغيرات تتعلق بالولايات المتحدة للفترة من الربع الأول من عام 1947 وحتى الربع الأخير من عام 1992. ولكي نكون أكثر دقة، فإن:

R= سعر أذونات الخزينة لفترة ثلاثة أشهر.

M= عرض النقود $(M2)$ بملايين الدولارات.

P= مستوى الأسعار مقاساً بمخفض إجمالي الناتج المحلي (رقم قياسي للأسعار حيث 1987=100).

Y= إجمالي الناتج المحلي الحقيقي ببلايين الدولارات (بأسعار عام 1987).

قبل البدء في إجراء التحليل باستخدام بيانات السلسلة الزمنية، يجب عليك التحقق من جذور الوحدة. تذكر أنه في حالة وجود جذور وحدة وغياب التكامل الذاتي، فسوف تحدث مشكلة الانحدار الزائف. وفي هذه الحالة يجب عليك استخدام فروقات البيانات. وكبديل في حالة وجود جذر وحدة ووجود تكامل ذاتي، فسوف تحصل على معلومات اقتصادية قيمة تؤكد أن للسلسلة اتجاهاً عاماً مشتركاً.

تشير الاختبارات في الحالة الراهنة، إلى أننا لانستطيع رفض فرضية وجود جذور وحدة في كافة المتغيرات مع غياب التكامل الذاتي. ولتجنب مشكلة الانحدار الزائف، سنستخدم فروقات البيانات. وبصفة خاصة، سنستخدم اللوغاريثمات لكل سلسلة ثم نأخذ الفروقات للسلسلة اللوغريثمية ونضربها في 100. وهذا يعنى ضمناً أننا نستخدم التغيرات النسبية في كل متغير (مثلا: قيمة 1 تعنى تغيراً مقداره 1%). وهكذا، فإن:

ΔR = النسبة المئوية للتغير في سعر الفائدة.

ΔM = النسبة المئوية لتغير عرض النقود.

ΔP = النسبة المئوية للتغير في مستوى الأسعار (التضخم).

ΔY = النسبة المئوية للتغير في إجمالي الناتج المحلي (نمو إجمالي الناتج المحلي).

تمرين رقم (11-5):

باستخدام البيانات حول R,M,P,Y في الملف RMPY.XLS:

أ) أجر اختبار التحقق من جذور الوحدة لكل متغير من المتغيرات.

ب) أجر اختبار التحقق من التكامل الذاتي بين المتغيرات باستخدام البيانات حول $\Delta R, \Delta M, \Delta P, \Delta Y$ في الملف RMPY.XLS:

ت) أجر اختبار التحقق من جذور الوحدة في كل متغير من المتغيرات.

جدول رقم (11- 4): نموذج $VAR(1)$ مع $\Delta R, \Delta M, \Delta P, \Delta Y$ كمتغيرات تابعة

	المتغير التابع ΔR		المتغير التابع ΔM		المتغير التابع ΔP		المتغير التابع ΔY	
	المعامل (Coeff)	قيمة P (P-val.)	المعامل (Coeff)	قيمة P (P-val.)	المعامل (Coeff)	قيمة P (P-val.)	المعامل (Coeff)	قيمة P (P-val.)
قاطع	-3.631	0.162	0.335	0.001	0.161	0.138	0.495	0.005
ΔR_{t-1}	0.222	0.003	-0.013	2.0E-5	0.010	0.002	3.8E-4	0.940
ΔM_{t-1}	3.391	0.007	0.749	1.0E-33	0.121	0.021	0.283	9.3E-4
ΔP_{t-1}	1.779	0.228	0.061	0.303	0.519	1.0E-14	-0.117	0.242
ΔY_{t-1}	3.224	0.004	-0.032	0.480	-0.039	0.407	0.309	7.0E-5
الوقت	-0.056	0.011	3.4E-4	0.695	0.002	0.048	-0.003	0.035

يوضح الجدول رقم (11-4) نتائج تقدير المربعات الصغرى الاعتيادية في نموذج $VAR(1)$. لاحظ أن هذا الجدول يختلف قليلاً من حيث الشكل عن الجداول السابقة. ونظراً لوجود أربعة متغيرات $(\Delta R, \Delta M, \Delta P, \Delta Y)$في نموذج VAR، فلابد من تقدير أربع معادلات. لقد وضعنا نتائج جميع المعادلات في جدول واحد.

تتناول كل معادلة انحدار متغير تابع على متباطئة واحدة لكافة متغيرات نموذج VAR. ولتوفير المساحة، أدرجنا تقدير مربعات صغرى اعتيادية واحد وقيمة P واحدة لكل معامل.

فإذا اختبرنا معنوية المعاملات (أى المعاملات التى تقل قيمة P فيها عن 0.05)، تظهر لنا بعض الأنماط المهمة.

أولاً: يعد لوغاريثم المتغير التابع في كل معادلة معنوياً. وهذا يعني أنه في المعادلة التي تتضمن ΔR_t كمتغير تابع، يوفر ΔR_{t-1} قدرة تفسيرية مهمة. وفي المعادلة التي تتضمن ΔM_t كمتغير تابع، يوفر ΔM_{t-1} قدرة تفسيرية مهمة،... الخ.

ثانياً: توضح نتائج المعادلات الأربعة أنماطاً مهمة لسببية جرانجر. ففي المعادلة التى تتضمن ΔR كمتغير تابع، يمكننا مشاهدة أن نمو إجمالي الناتج المحلي ونمو عرض النقود يتمتعان بقدرة تفسيرية للتغيرات الراهنة في سعر الفائدة. وفى حالة العلاقة بين سعر الفائدة وعرض النقود، $\frac{\Delta R}{\Delta M}$(سعر الفائدة ÷ عرض النقود)، تظهر المعادلة التى تتضمن ΔM كمتغير تابع، أن السببية تتم في الاتجاهين لأن التغير في سعر الفائدة يؤدي أيضاً إلى نمو عرض النقود حسب سببية جرانجر. ورغم ذلك، فإن التغير في سعر الفائدة لايسبب نمو إجمالي الناتج المحلي حسب سببية جرانجر. وتعد نتائج سببية جرانجر المتعلقة بالتضخم ذات أهمية خاصة حيث ينظر إلى التضخم على أنه يسبب أى متغير آخر حسب سببية جرانجر ولكن كلا من ΔR وΔM يسببان التضخم حسب سببية جرانجر.

بإمكان خبير الاقتصاد الكلى استخدام هذه النتائج لمعالجة المشكلات النظرية ذات الأهمية (مثلاً: هل التضخم ظاهرة نقدية بحتة؟ هل يتم دعم مرئيات الاقتصاديين النقديين المتعلقة بالاقتصاد؟ هل يتم دعم المرئيات الكينزية المتعلقة بالاقتصاد؟ هل يتأثر الاقتصاد الحقيقي بالتضخم؟ ... الخ)، ولكن ذلك يعد خارج نطاق البحث التفصيلي في هذا الكتاب.

اختيار قيمة المتباطئة في نماذج متجه الانحدار الذاتي (VAR):

تستند نتائج المثال السابق إلى نموذج $VAR(1)$. أى أننا حددنا $p=1$ واستخدمنا متباطئة واحدة لكل متغير لتفسير المتغير التابع. وبالطبع، وبصفة عامة، نحن في حاجة إلى تحديد قيم أخرى لـ p غير الواحد. يتوافر عدد كبير من الأدبيات التي تناقش موضوع اختيار قيمة المتباطئة، ولكن معظم المعايير المقترحة معقدة جداً بحيث يصعب حسابها من خلال برنامج مثل إكسل. فإذا كنت مهتماً بالتعامل مع نماذج VAR على نطاق واسع، فسوف يكون من الأفضل قراءة المزيد حول الموضوع وإجادة برنامج حاسب آلي خاص بالاقتصاد القياسي مثل MicroFit.

تتيح إحصائية t وقيم الاحتمال p المستخدمة في هذا الكتاب بأسره، معلومات قيمة حول قيمة المتباطئة. ويتضح ذلك جليا في المثال التالي:

مثال: متجه نموذج انحدار ذاتي $VAR(2)$ مع متغيرات RMPY:

لقد استخدمنا في المثال السابق بيانات عن $\Delta R, \Delta M, \Delta P, \Delta Y$ لتقدير نموذج $VAR(1)$.

يكرر الجدول رقم (11-5) التحليل باستخدام نموذج $VAR(2)$. إن العديد من المعاملات للمتغيرات قبل فترتين معنوية. فعلى سبيل المثال، تعد ΔR_{t-2} معنوية في المعادلة التى تتضمن ΔR_t كمتغير تابع. وهذا يشير إلى أن نموذج $VAR(1)$ المستخدم في المثال السابق لم يكن ملائماً.

لإعطاء فكرة عن تكاليف استخدام نموذج خاطئ، ألق نظرة على معادلة تتضمن ΔY كمتغير تابع. تذكر أننا في نموذج $VAR(1)$ استنتجنا أن التضخم لم

يسبب نمو إجمالي الناتج المحلي حسب سببية جرانجر. ولكن نموذج $VAR(2)$ يشير إلى أن التضخم يسبب نمو إجمالي الناتج المحلي حسب سببية جرانجر. ولأن العلاقة بين نمو إجمالي الناتج المحلي ومعدل التضخم هو مصدر الجدل الواسع في الاقتصاد الكلي الحديث، فإن تكلفة الاستخدام الخاطئ لنموذج $VAR(1)$ كبيرة جداً.

جدول رقم (5-11) نموذج رمبي $VAR(2)$ RMPY

باستخدام $\Delta R, \Delta M, \Delta P, \Delta Y$ كمتغيرات تابعة

	المتغير التابع ΔR		المتغير التابع ΔM		المتغير التابع ΔP		المتغير التابع ΔY	
	المعامل (Coeff)	قيمة P (P-val.)	المعامل (Coeff)	قيمة P (P-val.)	المعامل (Coeff)	قيمة P (P-val.)	المعامل (Coeff)	قيمة P (P-val.)
قاطع(Interecept)	-4.000	0.103	0.261	0.017	0.113	0.311	0.513	0.006
ΔR_{t-1}	0.315	1.9E-5	-0.017	3.6E-7	0.009	0.004	0.002	0.670
ΔM_{t-1}	2.824	0.106	0.615	8.E-15	0.086	0.280	0.310	0.019
ΔP_{t-1}	3.049	0.061	-0.020	0.785	0.366	1.6E-6	0.074	0.545
ΔY_{t-1}	3.696	4.6E-4	-0.051	0.270	-0.010	0.835	0.270	0.001
ΔR_{t-2}	-0.346	4.5E-6	0.003	0.298	-0.001	0.795	-0.010	0.085
ΔM_{t-2}	-2.201	0.213	0.157	0.045	0.025	0.755	-0.094	0.480
ΔP_{t-2}	1.164	0.457	0.095	0.170	0.282	1.0E-4	-0.233	0.049
ΔY_{t-2}	1.085	0.303	0.036	0.445	-0.046	0.334	0.153	0.054
الوقت (Time)	-0.045	0.029	-2.0E-4	0.798	0.001	0.209	-0.003	0.104

تمرين رقم (11-6):

باستخدام نتائج الجدول رقم (11-5)، ناقش سببية جرانجر بين كافة المتغيرات في النموذج.

أوضح المثال السابق أهمية الاختيار الصحيح لقيمة المتباطئة في نموذج $VAR(p)$. وفي ظل غياب المزيد من الدراسة لهذه القضية، أوصينا باستخدام طريقة الاختبار التتابعى التالية:

الخطوة الأولى: اختر أقصى قيمة ممكنة للمتباطئة $p_{\max}$ تراها مناسبة.

الخطوة الثانية: قدر نموذج $VAR(p_{\max})$. فإذا كان أي من المتغيرات المتباطئة عدد $p_{\max}$ فترة معنوياً، استخدم نموذج $VAR(p_{\max})$. وإلا فانتقل إلى الخطوة التالية.

الخطوة الثالثة: قدر نموذج $VAR(p_{\max}-1)$. فإذا كان أي من المتغيرات المتباطئة عدد $VAR(p_{\max}-1)$ فترة معنوياً، استخدم نموذج $VAR(p_{\max}-1)$. وإلا فانتقل إلى الخطوة التالية.

الخطوة الرابعة: قدر نموذج $VAR(p_{\max}-2)$. فإذا كان أي من المتغيرات المتباطئة عدد $VAR(p_{\max}-2)$ فترة معنوياً، استخدم نموذج $VAR(p_{\max}-2)$. وإلا فانتقل إلى الخطوة التالية، وحاول $(p_{\max}-3)$، ... الخ.

تمرين رقم (11-7):

باستخدام المتغيرات $\Delta R, \Delta M, \Delta P, \Delta Y$ من ملف البيانات RMPY.XLS:

أ) ابدأ بـ: $p_{\max} = 5$ واختر قيمة المتباطئة المناسبة لنموذج VAR.

ب) باستخدام نتيجة البند (أ)، ناقش سببية جرانجر بين المتغيرات $\Delta R, \Delta M, \Delta P, \Delta Y$.

تمرين رقم (11-8):

يتضمن الملف LONGGDP.XLS، كما نتذكر، بيانات سنوية عن نصيب الفرد من إجمالي الناتج المحلي الحقيقي لأربع من أكبر الدول الناطقة بالإنجليزية (الولايات المتحدة، المملكة المتحدة، كندا وأستراليا) للفترة 1870-1993. احسب الفروقات الأولى للحصول على سلسلة زمنية لنمو نصيب الفرد من إجمالي الناتج المحلى لكل دولة من الدول الأربعة. قدر نموذج VAR باستخدام هذه البيانات.

التوقع (التنبؤ) باستخدام نماذج متجه الانحدار الذاتي (VAR) :

لقد قلنا القليل فقط في هذا الكتاب حتى الآن عن التوقع على الرغم من حقيقة كونها تمثل نشاطاً مهما لدى الاقتصاديين. هناك سببان وراء حذف الموضوع.

أولاً: يعد مجال التوقع واسعاً. ففي ظل الكم الهائل من الأبحاث والمشكلات المطلوب بحثها، من المستحيل إعطاء الموضوع حقه في مثل هذا الكتاب[7].

ثانياً: إن إجراء التوقعات الأساسية باستخدام الحاسب الآلي إما سهل جداً أو صعب جداً وذلك اعتماداً على نوع البرنامج المستخدم. ولكي نكون أكثر دقة، يتمتع العديد من برامج الحاسب الآلي (مثلاً MicroFit) بخواص التوقعات سهلة الاستخدام. بمجرد تقدير النموذج (VAR أو AR مثلاً)، بإمكانك عمل التوقعات بسهولة من خلال النقر على الخيار المناسب. بعبارة أخرى، يتيح العديد من برامج الحاسب الآلي إمكانية عمل التوقعات الأساسية دون الحاجة إلى معرفة أعمق بالموضوع. ولكن برامج الجداول الإلكترونية، (مثل إكسل)، لاتتمتع بخواص إعداد التوقعات للنماذج المستخدمة في هذا الكتاب. بالإمكان حساب التوقعات ولكن ذلك غير ملائم نظراً للحاجة إلى كتابة معادلات طويلة ومعقدة.

وفي ضوء هذه المشكلات، سنعطي مقدمة موجزة فقط لبعض المشكلات العملية والأفكار البديهية المرتبطة بموضوع إجراء التوقعات. وسوف تقتصر كافة مناقشاتنا على إجراء التوقعات باستخدام نماذج VAR. ولكن من الضروري الإشارة إلى أن الأفكار ترتبط أيضاً بإجراء التوقعات بنماذج السلاسل الزمنية الأحادية. خلاصة القول، فإن نموذج AR ما هو إلا نموذج VAR بمعادلة واحدة.

يتم إجراء التوقعات عادة باستخدام متغيرات السلاسل الزمنية، حيث تتمثل الفكرة في استخدام البيانات المشاهدة للمتوقع حدوثه في المستقبل. وبأسلوب فني أكثر، يتم استخدام بيانات للفترات t=1,....,T للتنبؤ بالفترات T+1,T+2,....,etc.

لمزيد من التوضيح حول كيفية إعداد التوقعات، تناول نموذج $VAR(1)$ بمتغيرين Y و X:

$$Y_t = \alpha_1 + \delta_1 t + \phi_{11} Y_{t-1} + \beta_{11} X_{t-1} + e_{1t},$$

$$X_t = \alpha_2 + \delta_2 t + \phi_{21} Y_{t-1} + \beta_{21} X_{t-1} + e_{21t}.$$

ليس بإمكانك مشاهدة Y_{T+1} ولكنك تريد تخمين ماهو محتمل. باستخدام المعادلة الأولى لنموذج VAR وتحديد t=T+1، سنحصل على عبارة جبرية تخص Y_{T+1}:

$$Y_{T+1} = \alpha_1 + \delta_1(T+1) + \phi_{11}Y_T + \beta_{11}X_T + e_{1T+1}.$$

لايمكن استخدام هذه المعادلة مباشرة للحصول على Y_{T+1} لأننا لا نعرف قيمة e_{1T+1}. أى إننا لانعرف ماهى الصدمة أو المفاجأة غير المتوقعة التى ستحصل في الاقتصاد في الفترة القادمة. كما أننا أيضا لا نعرف ماهى المعاملات. ولكن إذا تجاهلنا حد الخطأ (الذي لايمكن التوقع به) واستبدلنا المعاملات بتقديرات المربعات الصغرى الاعتيادية، فسوف نحصل على التوقع الذي يحمل الرمز $\hat{Y}_{T+1}$:

$$\hat{Y}_{T+1} = \hat{\alpha}_1 + \hat{\delta}_1(T+1) + \hat{\phi}_{11}Y_T + \hat{\beta}_{11}X_T.$$

فإذا كنت تستخدم برنامج جداول إلكترونية مثل إكسل، لاحظ أن كل مايخص $\hat{Y}_{T+1}$ في المعادلة يمكن أخذها إما من البيانات الأصلية أو من مخرجات أمر الانحدار. من السهل نظرياً إدخال كافة الأرقام المفردة (أى تقديرات المربعات الصغرى الاعتيادية للمعاملات $(Y_T, X_T, T+1)$ في المعادلة لحساب $\hat{Y}_{T+1}$.
كما يمكن استخدام طريقة مماثلة للحصول على $\hat{X}_{T+1}$. وسوف تعرف كيف أن الحساب العملي لهذه التوقعات بهذه الكيفية قد يكون غير ملائم ويستغرق وقتاً طويلاً. لذلك إذا كنت تخطط لإجراء المزيد من التوقعات، يفضل تجنب برامج

الحاسب مثل إكسل واستخدام برامج الحاسب الآلي المتخصصة في مجال الاقتصاد القياسي مثل MicroFit.

تصف الفقرة السابقة كيفية التوقع بفترة واحدة في المستقبل. ولكن بالإمكان استخدام الطريقة ذاتها لفترتين شريطة عمل توسعة إضافية واحدة للنموذج. ففي حالة الفترة الواحدة استخدمنا X_T و Y_T للحصول على $\hat{Y}_{T+1}$ و $\hat{X}_{T+1}$. أما في حالة الفترتين فإن $\hat{X}_{T+2}$ و $\hat{Y}_{T+2}$ يعتمدان على $\hat{X}_{T+1}$ و $\hat{Y}_{T+1}$. ولكن بما أن بياناتنا تسري حتى الفترة T فإننا لانعرف ماهى X_{T+1} و Y_{T+1}. لذلك سنقوم بإحلالهما بـ $\hat{X}_{T+1}$ و $\hat{Y}_{T+1}$ أى إننا سنستخدم المعادلة الملائمة من نموذج VAR ونتجاهل الخطأ ونستبدل المعاملات بتقديرات المربعات الصغرى الاعتيادية ونستبدل القيم السابقة للمتغيرات التى لانشاهدها بتوقعاتها في المعادلة:

$$\hat{Y}_{T+2} = \hat{\alpha}_1 + \hat{\delta}_1 (T+2) + \hat{\phi}_{11} \hat{Y}_{T+1} + \hat{\beta}_{11} \hat{X}_{T+1}.$$

يمكن حساب المعادلة أعلاه باستخدام برنامج حاسب آلي مثل إكسل على الرغم من عدم الملاءمة إلى حد ما. ويمكن حساب $\hat{X}_{T+2}$ بطريقة مماثلة باستخدام المعادلة التالية:

$$\hat{X}_{T+2} = \hat{\alpha}_2 + \hat{\delta}_2 (T+2) + \hat{\phi}_{21} \hat{Y}_{T+1} + \hat{\beta}_{21} \hat{X}_{T+1}.$$

بإمكاننا استخدام الطريقة العامة لاستبعاد الخطأ، استبدال المعاملات بتقديرات المربعات الصغرى الاعتيادية واستبدال القيم المتباطئة للمتغيرات غير المشاهدة بالتوقعات، بهدف الحصول على توقعات لأي عدد من الفترات المستقبلية لأي نموذج $VAR(p)$.

أوضح النقاش السابق كيفية حساب تقديرات النقاط الأساسية للتوقعات. وعملياً، فإنه نادراً ما يتطابق ما يحدث مع توقعك. لقد ناقشنا هذه المشكلة في الفصل الخامس إذ أشرنا إلى أن المربعات الصغرى الاعتيادية تعطي تقديرات فقط للمعاملات وأنها لن تكون صحيحة بدرجة كبيرة. لذلك، نوصي أيضاً، بالإضافة إلى تقديرات المربعات الصغرى الاعتيادية، إدخال فترات ثقة لعكس مستوى عدم التأكد بشأن تقدير المعامل. عند إجراء التوقعات، يمكن حساب فترات الثقة التي قد توفر معلومات قيمة. فعلى سبيل المثال، أصبح من المألوف في الجهات الحكومية تقديم فترات ثقة لتوقعاتها. إن بنك إنجلترا، مثلاً، يصيغ أحياناً عبارات مثل: "إن توقعنا لمستوى التضخم في العام القادم هو 1.8% وأننا واثقون بنسبة 95% أن مستوى التضخم سيكون في المدى 1.45%-2.15%". ويوفر العديد من برامج الحاسب الآلي فترات الثقة تلقائياً ومن ثم لا داعي لمعرفة المعادلة الدقيقة عند إجراء التوقعات. فإذا كنت تستخدم برنامج حاسب آلي مثل أكسل، فإن المعادلة تعد معقدة جداً ومن ثم لا يمكن تقديرها؛ وهذا هو السبب وراء عدم عرضها هنا.

مثال: نموذج متجه انحدار ذاتي VAR(2) مع متغيرات RMPY:

سنتناول في هذا المثال إجراء التوقعات باستخدام بيانات عن $\Delta R, \Delta M, \Delta P, \Delta Y$ من الملف RMPY.XLS. وكما الحال أعلاه، سيتم استخدام $VAR(2)$ علماً بأن لدينا بيانات عن المتغيرات من الربع الثاني لعام 1947 حتى الربع الرابع لعام 1992. إن الممارسة المألوفة هي الإبقاء على بعض البيانات بغرض مقارنة التوقعات معها. وهنا سنستخدم بيانات من الربع الثاني لعام 1947 وحتى الربع الرابع لعام 1991 لتقدير نموذج VAR(2). سنجري التوقعات من الربع الأول لعام 1992 وحتى الربع الرابع لعام 1992 ثم نقارن توقعاتنا لعام 1992 مع ما حدث فعلاً عام 1992.

يحتوي جدول رقم (11-6) التوقعات والمشاهدات الفعلية للتضخم ونمو إجمالي الناتج المحلي لعام 1992 (بعكس كافة الأمثلة العملية الأخرى في هذا الكتاب، تم حساب هذه التوقعات باستخدام برنامج MicroFit وليس برنامج إكسل).

جدول رقم (11-6): توقعات التضخم ونمو إجمالي الناتج المحلي لعام 1992 باستخدام نموذج $VAR(2)$ مع متغيرات RMPY

	ΔP		ΔY	
	المتوقع (Forecast)	الفعلي (Actual)	المتوقع (Forecast)	الفعلي (Actual)
1992Q1	0.626	0.929	-0.019	0.865
1992Q2	0.731	0.689	0.220	0.698
1992Q3	0.862	0.289	0.275	0.838
1992Q4	0.940	0.813	0.271	1.393

لتوضيح ذلك، لاحظ أن كافة الأرقام عبارة عن تغيرات نسبية لربع السنة. فعلى سبيل المثال، بلغ معدل التضخم المتوقع للربع الثاني لعام 1992 مقدار 0.731، أي معدل تضخم سنوي مقداره 2.96%. يتضح من الجدول أن نموذج $VAR(2)$ كان جيداً من حيث التوقع بمعدلات التضخم باستثناء الربع الثالث من عام 1992 عندما كان معدل التضخم منخفضاً بصورة غير عادية. ولكن الأداء بالنسبة لتوقعات نمو إجمالي الناتج المحلي لم يكن بالجودة ذاتها حيث توقع نموذج VAR بمعدلات نمو أقل من المعدلات الفعلية.

تمرين رقم (11-9):

يوصى بحل هذا التمرين وغيره من التمارين التي تتضمن إجراء التوقعات فقط إذا كان لديك برنامج حاسب آلي خاص بالاقتصاد القياسي قادراً على إجراء التوقعات. فإذا كنت تتعامل مع برنامج حاسب آلي مثل إكسل فإن حل هذا التمرين سيكون صعباً.

استخدم المتغيرات $\Delta R, \Delta M, \Delta P, \Delta Y$ من الملف RMPY.XLS .

أ- لقد استخدمنا نموذج $VAR(2)$ في المثال السابق. باستخدام نموذج $VAR(p)$ لقيم p المختلفة (مثلاً $p = 1,3,4,...$) أجر توقعات لعام 1992. هل يحقق أي من نماذج VAR التي استخدمتها توقعات أفضل من تلك الموضحة في جدول رقم (11-6)؟

ب- لقد استخدمنا في المثال السابق بيانات للفترة من الربع الثاني لعام 1947 حتى الربع الرابع لعام 1992، لتقدير نموذج VAR الذي استخدم بدوره لإجراء توقعات للفترة من الربع الأول لعام 1992 وحتى الربع الرابع من عام 1992. أعد المثال أعلاه باستخدام بيانات للفترة من الربع الثاني لعام 1947 وحتى الربع الرابع لعام 1990 لإجراء توقعات للفترة من الربع الأول لعام 1991 وحتى الربع الرابع لعام 1992 (أي توقعات لعامين بدلاً عن واحد).

ت- حاول إجراء توقعات لفترات أطول. مثلاً لقد تم في البند (ب) إعداد توقعات لفترة سنتين. حاول الآن إعداد توقعات لفترة ثلاث سنوات، أربع سنوات، خمس سنوات،... الخ. فسر نتائجك. هل تشير نتائجك إلى أن نماذج VAR قادرة على إعداد التوقعات قصيرة المدى على نحو أفضل من التوقعات بعيدة المدى؟

تمرين رقم (11-10):

باستخدام بيانات تمرين رقم (11-8) ونموذج VAR الذي تم إنشاؤه هناك، قم بإجراء توقعات لنمو إجمالي الناتج المحلي للدول الأربع المذكورة في المثال. حاول تجربة فترات مختلفة للتوقعات. هل كان نموذج VAR جيداً في إجراء التوقعات؟

متجه الانحدار الذاتي مع متغيرات متكاملة ذاتياً:

في الشرح السابق لنماذج VAR افترضنا أن كافة المتغيرات مستقرة. فإذا كانت بعض المتغيرات الأصلية ذات جذور وحدة وغير متكاملة ذاتياً، فإنه يجب أخذ فروقات للمتغيرات التى تحتوى جذور وحدة واستخدام المتغيرات الناتجة المستقرة في نموذج VAR. وهذا ينطبق على كل حالة باستثناء تلك التي تكون فيها المتغيرات ذات جذور وحدة وتكون متكاملة ذاتياً.

تذكر أنه في مثل هذه الحالة المتعلقة بمناقشة سببية جرانجر، أوصينا باستخدام نموذج تصحيح الخطأ (ECM). بالإمكان استخدام الطريقة ذاتها هنا. وبدلاً عن التعامل مع نموذج متجه الانحدار الذاتي، يمكنك بصفة خاصة التعامل مع نموذج "متجه تصحيح الخطأ" (VECM). وكما الحال بالنسبة لنموذج VAR ستكون لنموذج VECM معادلة واحدة لكل متغير في النموذج. وفي حالة متغيرين Y و X، سيكون نموذج VECM على النحو التالي:

$$\Delta Y_t = \varphi_1 + \delta_1 t + \lambda_1 e_{t-1} + \gamma_{11}\Delta Y_{t-1} + \ldots + \gamma_{1p}\Delta Y_{t-p} + \omega_{11}\Delta X_{t-1} + \ldots + \omega_{1q}\Delta X_{t-q} + \in_{1t} .$$

$$\Delta X_t = \varphi_2 + \delta_2 t + \lambda_2 e_{t-1} + \gamma_{21}\Delta Y_{t-1} + \ldots + \gamma_{2p}\Delta Y_{t-p} + \omega_{21}\Delta X_{t-1} + \ldots + \omega_{2q}\Delta X_{t-q} + \in_{2t} .$$

وكما سبق، فإن $e_{t-1}=Y_{t-1}-\alpha-\beta X_{t-1}$. لاحظ أن نموذج VECM ما هو إلا نموذج VAR مع متغيرات خاضعة للفروقات باستثناء الحد e_{t-1}. بالإمكان الحصول على تقدير متغير تصحيح الخطأ بإجراء انحدار المربعات الصغرى الاعتيادية Y على X وحفظ البواقي. بعد ذلك يمكن استخدام المربعات الصغرى الاعتيادية لتقدير نماذج ECM والحصول على قيم P وفترات الثقة. كما يمكن اختبار قيمة المتباطئة وإجراء التوقعات بطريقة مماثلة لنموذج VAR مع قدر ضئيل من التعقيد الإضافي المتمثل في أن حد تصحيح الخطأ، e_t، يجب حسابه. ولكن الأمر سهل باستخدام تقديرات المربعات الصغرى الاعتيادية لـ α وβ واستبدال حد الخطأ الباقي u_t.

تمرين رقم (11-11):

بالنسبة لهذا التمرين، استخدم بيانات أسعار البرتقال العادي والبرتقال العضوي المضمنة في الملف ORANGE.XLS.

أ) بدءاً بـ: $p_{\max}=4$، اختر قيمة متباطئة لنموذج VECM وقدر كل معادلة باستخدام المربعات الصغرى الاعتيادية.

ب) باستخدام نموذج VECM من البند (أ)، أجر عملية إعداد التوقعات. حاول تجربة فترات زمنية مختلفة للتوقعات. هل يحقق نموذج VAR توقعات جيدة؟

تمرين رقم (11-12):

استخدم بيانات عن: Y= الاستهلاك، X= الدخل الشخصي وذلك من الملف :INCOME.XLS

أ) افترض (ربما بصورة غير صحيحة في ضوء تمرين رقم 10-5) أن كلا من YوX متكاملة ذاتياً. كرر خطوات تمرين رقم (11-11) لإجراء توقعات.

ب) افترض الآن أن YوX تحتويان على جذور وحدة ولكنهما غير متكاملتين ذاتياً. أنشيء نموذج VAR باستخدام بيانات الفروقات (أى ΔY وΔX) وأجر عملية إعداد التوقعات.

ت) قارن نتائج البندين أ،ب. ما هو التأثير المتوقع عن افتراض (ربما بصورة غير صحيحة) وجود التكامل على أداء التوقعات؟

ملخص الفصل:

1- إن العديد من متغيرات السلاسل الزمنية، وخاصة أسعار الأصول المالية (الأوراق المالية)، تبدو وكأنها تظهر سلوك مسارات عشوائية. لهذا، من الصعب التوقع بكيفية تغيرها في المستقبل. ولكن غالباً ما تعكس هذه المتغيرات أنماط تقلب قابلة للتوقع.

2- يعد مربع التغير في سعر الأصل مقياساً للتقلب.

3- بالإمكان استخدام أساليب السلاسل الزمنية القياسية لنمذجة أنماط التقلبات في أسعار الأصول. ولكن الفرق الوحيد هو أن تقلب سعر الأصل هو المستخدم في هذه الحالة وليس سعر الأصل ذاته.

4- يكون X سبباً في حدوث Y حسب سببية جرانجر إذا كانت قيم X السابقة تتمتع بقدرة تفسيرية لـ Y.

5- إذا كانت YوX مستقرتين، فبالإمكان استخدام الأساليب الإحصائية القياسية المستندة إلى نموذج ADL لاختبار سببية جرانجر.

6- إذا كانت YوX تحتويان على جذور وحدة ومتكاملة ذاتياً، فبالإمكان استخدام الأساليب الإحصائية المستندة إلى نموذج تصحيح الخطأ ECM لاختبار سببية جرانجر.

7- لنموذج متجه الانحدار الذاتي معادلة واحدة لكل متغير قيد الدراسة. وتختار كل معادلة متغيراً واحداً كمتغير تابع. وتكون المتغيرات التفسيرية عبارة عن متباطئات لجميع المتغيرات محل الدراسة.

8- تعد نماذج VAR مهمة جداً لإجراء التوقعات، وكذلك لاختبار سببية جرانجر، أو بصفة عامة، مهمة لفهم العلاقات بين عدة سلاسل بيانات.

9- إذا كانت جميع متغيرات نموذج VAR مستقرة، فبالإمكان استخدام المربعات الصغرى الاعتيادية لتقدير كل معادلة واستخدام الأساليب الإحصائية القياسية (مثلاً، يمكن استخدام قيم الاحتمال p وإحصائية t لاختبار معنوية المتغيرات).

10- إذا كانت المتغيرات قيد الدراسة ذات جذور وحدة ومتكاملة ذاتياً، فيجب استخدام نسخة VAR التى تسمي "نموذج متجه تصحيح الخطأ" (VECM).

ملحق رقم (11-1): اختبارات الفروض لأكثر من معامل:

ناقشنا، في الفصلين الخامس والسادس، إحصائية F التي تستخدم في اختبار فرضية أن $R^2 = 0$ في نموذج الانحدار المتعدد:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e.$$

لقد أوضحنا كيف أن هذه المعادلة مساوية لاختبار فرض العدم: $H_0: \beta_1 = ... = \beta_k = 0$ (أى معرفة ما إذا كانت كافة معاملات الانحدار مجتمعة تساوي الصفر). كما ناقشنا اختبار معنوية المعاملات المفردة باستخدام إحصائية t و قيمة الاحتمال p.

ولكن ليس لدينا أدوات لاختبار الحالات الوسيطة (مثلا: في حالة k=4، قد يكون اهتمامنا منصباً على اختبار فرض العدم: ($H_0: \beta_1 = \beta_2 = 0$). لقد برزت مثل هذه الحالات في مناقشتنا لسببية جرانجر (مثلاً: كان لدينا نموذج انحدار بأربع فترات تباطؤ لتضخم الأسعار وأربع فترات تباطؤ لتضخم الأجور واتجاهاً عاماً غير عشوائي وكنا مهتمين باختبار ما إذا كانت جميع معاملات المتباطئات الأربعة لتضخم الأجور مساوية للصفر). إن الهدف من هذا الملحق هو وصف إجراءات وقاعدة تقريبية لإجراء مثل هذه الأنواع من الاختبارات.

تتم الإشارة إلى إحصائية F، الموضحة في الفصل الخامس، بهذا الاسم لأنها واحدة فقط من فئة واسعة من إحصائيات الاختبارات التي تستمد قيمها الحرجة من الجداول الإحصائية لتوزيع F. لقد تناولنا في هذا الكتاب، حتى الآن، قدراً ضئيلاً من النظرية الإحصائية ولم نوضح كيفية استخدام الجداول الإحصائية. ولكن إذا كنت تخطط للتعامل بشكل واسع مع اختبار سببية جرانجر، فإننى

أنصحك بدراسة مرجع في أساسيات علم الإحصاء أو الاقتصاد القياسي لمعرفة المزيد عن الأسس الإحصائية لاختبار الفروض.

لفهم الإجراء الأساسي لاختبار F، سنوضح الفرق بين نماذج الانحدار المقيدة (restricted) وغير المقيدة (unrestricted). أى أن معظم الفرضيات التى تود اختبارها تضع قيوداً على النموذج. لذلك يمكننا التمييز بين الانحدار الذي يتم فرض قيود عليه والانحدار الذي لا يتضمن أي قيود. فإذا كان نموذج الانحدار الخالي من القيود، مثلاً، هو:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + e.$$

وترغب اختبار فرض العدم $H_0: \beta_2 = \beta_4 = 0$ ، فإن نموذج الانحدار المقيد سيكون على النحو التالي:

$$Y = \alpha + \beta_1 X_1 + \beta_3 X_3 + e.$$

تتمثل الطريقة العامة لاختبار الفروض في أنه يتم أولاً حساب إحصائية الاختبار ومقارنته مع قيمة حرجة. فإذا كانت إحصائية الاختبار أكبر من القيمة الحرجة (من حيث القيمة المطلقة) يتم رفض الفرضية، وإلا يتم قبولها. بإيجاز، هناك دائماً عنصران لإجراء اختبار الفروض: إحصائية اختبار وقيمة حرجة.

تسمى إحصائية الاختبار هنا عادة بإحصائية F وتكون على النحو التالي:

$$f = \frac{(R_U^2 - R_R^2)/J}{R_U^2/(T-k)},$$

حيث إن: R_U^2 و R_R^2 عبارة عن R^2 من نماذج الانحدار غير المقيدة والمقيدة على التوالي وأن J هو عدد القيود (مثلاً: J=2 في مثالنا نظراً لأن

$\beta_2 = 0$ و $\beta_4 = 0$ يمثلا قيدين). كما أن T هو عدد المشاهدات و k عدد المتغيرات التفسيرية في الانحدار غير المقيد.

لاحظ أن إحصائية F يمكن الحصول عليها بإجراء انحدارات غير مقيدة وانحدارات مقيدة (مثلاً: إجراء انحدار Y على X_4 و X_3 و X_2 و X_1 للحصول على R_U^2، ثم إجراء انحدار Y على X_3 و X_1 للحصول على R_R^2) ثم حساب المعادلة أعلاه باستخدام برنامج مثل اكسل أو آلة حاسبة. بإمكان بعض البرامج الاقتصادية القياسية المتخصصة (MicroFit مثلاً) حساب إحصائية F تلقائياً إذا حددت الفرضية قيد الاختبار .

إن الإجراء الخاص بالحصول على القيمة الحرجة لمقارنة إحصائية F معها، ينطوي على قدر من المشكلات. فالقيمة الحرجة تعتمد منهجياً على T-k وJ. يتضمن الجدول (11-7) القيم الحرجة التي يمكن استخدامها كقاعدة تقريبية إذا كان T-k كبيراً.

مثلاً، إذا كان لديك عدد كبير من المشاهدات لاختبار J=2 (أي قيدين: $\beta_2 = 0$ و $\beta_4 = 0$) وأردت استخدام مستوى معنوية 5%، فسوف تستخدم قيمة حرجة مقدارها 3.00 لمقارنة إحصائية F معها.

للتوضيح أكثر، لاحظ أن حالة J=1 لم يتم إدراجها نظراً لأن اختبار قيد واحد تقوم به إحصائية t أصلاً. لاحظ أيضاً أن القيم الحرجة تصغر كلما كبر عدد القيود. ويمكن استخدام هذه الحقيقة لتقريب القيم الحرجة لقيم J غير المدرجة في الجدول رقم (11-7) .

فعلى سبيل المثال، ستقع القيمة الحرجة لاختبار فرضية تحتوى على 7 قيود (J=7) بين القيمتين الحرجتين J=5 و J=10 في الجدول رقم (11-7). وفي

معظم الحالات، ستكون معرفة أن القيمة الحرجة الصحيحة تقع بين قيمتين كافية لتقرير قبول الفرضية أو رفضها. بناءً على ذلك، وعلى الرغم من أن جدول رقم (11-7) لا يتضمن كل قيم J الممكنة، إلا أنه يمكنك استخدامه إذا كان J مختلفاً عن تلك المذكورة أعلاه.

تعد الأرقام الحرجة في الجدول السابق صحيحة من الناحية المنهجية إذا كان $T-k=\infty$. فالقيم الحرجة لـ T- k>100 قريبة جداً من ذلك. ولإعطاء فكرة عن درجة سوء الخطأ إذا كان T-k<100، يوضح الجدول رقم (11-8) القيم الحرجة الصحيحة في حالة T- k=40.

يلاحظ أن جميع هذه القيم الحرجة أكبر قليلاً عن تلك الموضحة في الجدول بالنسبة لـ$T-k=\infty$. وقد تود استخدام هذه القيم إذا كان T- k =40 تقريباً.

جدول رقم (11-7) القيم الحرجة لاختبار F إذا كان (T-k) كبيراً

مستوى المعنوية (Significance Level)	J=2	J=3	J=4	J=5	J=10	J=20
5%	3.00	2.60	2.37	2.21	1.83	1.57
1%	4.61	3.78	3.32	3.02	2.32	1.88

جدول رقم (11-8) القيم الحرجة لاختبار F إذا كان (T-k=40)

مستوى المعنوية (Significance Level)	J=2	J=3	J=4	J=5	J=10	J=20
5%	3.23	2.92	2.69	2.53	2.08	1.84
1%	5.18	4.31	3.83	3.51	2.80	2.37

ورغم ذلك، سنقوم بإيضاحها هنا للحصول على فكرة عن الخطأ الذي قد يحدث نتيجة استخدام العينة الكبيرة للقيم الحرجة. فعلى سبيل المثال، إذا كان J=2 و T-k=40 وحصلت على إحصائية F مقدارها 4، فإن استخدام أي من الجدولين سيكون جيداً؛ فكليهما ينص على أن الفرضية يجب رفضها عند مستوى معنوية 5%. ورغم ذلك، إذا كانت إحصائية F مقدارها 3.1 فإنك سترفضها بصورة غير صحيحة باستخدام جدول العينة الكبيرة.

وإجمالاً: بإمكانك استخدام الأساليب والجداول الموضحة في الملحق بأمان تام في الحالات التالية:

1- إذا كان حجم العينة كبيراً بالمقارنة مع عدد المتغيرات التفسيرية (مثلاً T- k>100)، فإن جدول العينة الكبيرة أعلاه سيكون جيداً.

2- إذا كان T-k يساوي 40 تقريباً فإن جدول T-k=40 سيكون خياراً مضموناً.

3- إذا لم يكن T- k كبيراً ولا يساوي 40 تقريباً، فإن استخدام جدول T- k=40 سيكون مضموناً شرط أن لا تكون إحصائية الاختبار مقاربة للقيمة الحرجة وأن لا يكون T- k صغيراً جداً(مثلاً T-k<10).

خلاصة القول، ما دام لديك عدد كبير من المشاهدات أو أن بياناتك لا تقع في واحدة من هذه الحالات "المبهمة"، فلن تكون على قدر كبير من الخطأ عند استخدامك الأساليب الموضحة في هذا الملحق.

مثال: سببية جرانجر مع بيانات تضخم الأسعار وتضخم الأجور:

لقد أجرينا في الجزء الرئيسي من هذا الفصل، اختبارات سببية جرانجر باستخدام بيانات تضخم الأسعار وتضخم الأجور ووجدنا أن تضخم الأجور لايسبب تضخم الأسعار حسب سببية جرانجر، ولكن تضخم الأسعار

يسبب تضخم الأجور حسب سببية جرانجر. وسوف نحاول هنا التحقق عما إذا كانت هذه الاستنتاجات لا تزال صحيحة من خلال اختبارات F الصحيحة لسببية جرانجر.

تحقق أولاً عما إذا كان تضخم الأجور يسبب تضخم الأسعار حسب سببية جرانجر. لقد استخدمنا في الجزء الرئيسي من الفصل النموذج التالي غير المقيد حيث إن Y= تضخم الأسعار وX= تضخم الأجور:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \ldots + \phi_4 Y_{t-4} + \beta_1 X_{t-1} + \ldots + \beta_4 X_{t-4} + e_t$$

في هذا النموذج: $\hat{T} = 128$ [8] و $k = 9$ (أي p=q=4 بالإضافة إلى الاتجاه العام غير العشوائي في النموذج). ينتج عن تقدير المربعات الصغرى الاعتيادية R_U^2 مقداره 0.616.

تتمثل فرضية "عدم وجود سببية جرانجر" في: $H_0 = \beta_1 = \ldots = \beta_4 = 0$، والتي تتضمن أربعة قيود، ومن ثم فإن J=4. وفيما يلي نموذج الانحدار المقيد:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \ldots + \phi_4 X_{t-4} + e_t.$$

وينتج عن تقدير المربعات الصغرى الاعتيادية لهذا النموذج R_R^2 مقداره 0.613.

وباستخدام هذه الأرقام نجد أن قيمة إحصائية F هي 0.145.

وبما أن T- K=119 (كبير)، بإمكاننا مقارنة 0.145 مع القيمة الحرجة البالغة 2.37. وحيث إن 2.37 <0.145، لا نستطيع رفض الفرضية عند مستوى معنوية 5%. لذلك نقبل فرضية أن تضخم الأجور لا يسبب تضخم الأسعار حسب سببية جرانجر.

ولإجراء اختبار للتحقق عما إذا كان تضخم الأسعار يسبب تضخم الأجور، سنكرر الخطوات أعلاه باستثناء أن تضخم الأجور الآن يمثل المتغير التابع وتضخم الأسعار يمثل المتغير التفسيري. فإذا استخدمنا المربعات الصغرى الاعتيادية لتقدير الانحدار المقيد والانحدار غير المقيد، سنجد أن $R_U^2 = 0.605$ و $R_R^2 = 0.532$. لاحظ أن العناصر الأخرى في المعادلة بالنسبة لإحصائية F لا تتغير. وبإدخال هذه الأرقام في المعادلة بالنسبة لإحصائية F، نجد أن $f = 33.412$، وهو أكبر من القيمتين الحرجتين 1% أو 5%. وفي هذه الحالة، يمكنا رفض فرضية أن $\beta_1 = ... = \beta_4 = 0$ ونخلص إلى أن تضخم الأسعار يسبب تضخم الأجور وفقاً لسببية جرانجر.

لاحظ أن النتائج تقول بأن تضخم الأجور لا يسبب تضخم الأسعار وفقاً لسببية جرانجر ولكن تضخم الأسعار يسبب تضخم الأجور حسب سببية جرانجر، هي النتائج ذاتها التي تم التوصل إليها سابقاً في هذا الفصل.

ملاحظات ختامية:

1- بالطبع فإن هذا النموذج ليس هو النموذج الوحيد لسلوك أسعار الأسهم. وبعض هذه النماذج لاتعنى ضمناً وجود سلوك تحرك عشوائي. وهذا ليس هو المكان المناسب لتقديم شرح متكامل لأدبيات "كفاءة السوق" (market efficiency). ولكن يكفي الإشارة هنا إلى أن فرضية المسار العشوائي تعد شائعة جداً وقد حققت نجاحاً عملياً كبيراً في العديد من السياقات.

2- عموماً، يحتفظ المستثمرون بالأسهم أملاً في تحقيق أرباح منها. فهم يحققون أرباحاً من الأسهم في حالة ارتفاع قيمتها أو حصولهم على أرباح موزعة. لذلك يعد الارتفاع المتوقع في قيمة الأسهم عادة سبباً أساسياً لاحتفاظ المستثمرين بأسهمهم. وإذا لم تكن هذه الأسباب متاحة، لكان المستثمرون قد وضعوا أموالهم في أصول أكثر أماناً (السندات مثلا).

3- إن الاقتصادى المالي القادر على التوقع الدقيق لأسعار الأسهم الأكثر احتمالاً للارتفاع هو الشخص الأكثر احتمالاً لتحقيق ثراء بدلا عن إهدار وقته في كتابة تقرير قياسي عن أسعار الأسهم.

4- عند استخلاصنا لهذه النتيجة تجاهلنا الحد N-1 في مقام المعادلة الموضحة في الفصل الثاني. يجب عليك ملاحظة أن ذلك ليس مهما هنا. وفي بعض معادلات التباين يتم إحلال N محل N-1. وهنا N=1، لذلك يمكننا تجاهله.

5- لاحظ أن المتغير X_t قد تم حذفه من نموذج ADL(p,q) هذا نظراً لأن اختبارات سببية جرانجر تهدف لتحديد ما إذا كانت قيم X السابقة (وليس الحالية) قادرة على تفسير Y. فإذا أدرجنا X_t فسوف نكون قد سمحنا بالسببية بين القيم الحالية ولحدثت كافة الصعوبات التى أشير إليها سابقاً في هذا الكتاب بشأن تفسير العلاقات والانحدارات على أنها تعكس السببية. وربما تستغرب أيضاً من السبب الذي يجعلنا نستخدم نموذج ADL(p,q) كمقابل للنسخة التى يكون فيها ΔY_t هو المتغير التابع (انظر الفصل العاشر). إن السبب هو أنه من الأسهل تفسير سببية جرانجر في نموذج ADL(p,q) الأساسي هذا على أنها تعنى أن

المعاملات مساوية للصفر. لقد كان بإمكاننا تغطية موضوع هذا الجزء بأكمله باستخدام نموذجنا السابق، ADL(p,q)، ولكن ذلك سيؤدي إلى إجراء اختبارات فروض أكثر تعقيداً.

6- ثمة طريقة أكثر منهجية للتعبير عن المفاهيم في هذا الجزء وهى أن نماذج متجه الانحدار الذاتي (VARs) لا تعاني مشكلات التزامن (Simultaneity). وهذا هو الموضوع الذي سنتناوله بإيجاز في الفصل الثاني عشر تحت عنوان "المشكلات التي تستدعى استخدام نماذج الانحدار المتعدد".

7- أحد الكتب التمهيدية كتاب "نماذج السلاسل الزمنية لاقتصاديات الأعمال والتوقعات".

8- Time Series Models for Business Economics and Forecasting, by Philip Hans Frances (Cambridge University Press,1988).

9- تذكر أن إجراء أخذ المتغيرات وإدراج المتغيرات المتباطئة في الانحدار يقلل عدد المشاهدات. وهذا هو السبب وراء كون (T=128) بدلاً عن (T=133).

الفصل الثاني عشر

القيود والتوسعات
(Limitations and Extensions)

يعد الانحدار والأساليب الإحصائية المرتبطة به فعالة للغاية لملاءمتها للحل العملي لمجال واسع من المشكلات الاقتصادية المختلفة. ومع ذلك، فإذا تحققت الملاءمة التامة لحل كل مشكلة من تلك المشكلات، فإن خبراء الاقتصاد القياسي والإحصاء الساعين نحو تطوير اختبارات ونماذج ومقدرات جديدة، سيجدون أنفسهم عاطلين عن العمل. يقدم هذا الفصل وصفاً موجزاً لبعض قيود أساليب انحدار المربعات الصغرى الاعتيادية وبعض التوسعات والإضافات الملائمة المصممة للتعامل مع هذه القيود. وكما سنرى، فإن بعض الحالات لاتعني في أغلب الأحيان أن أسلوب المربعات الصغرى الاعتيادية هي الأسلوب الخاطئ بقدر ما تعني أن هناك مقدرات أفضل ينبغي استخدامها. وفى حالات أخرى، يعد استخدام طريقة المربعات الصغرى الاعتيادية غير مناسبة ويجب تجنبها.

من الضروري التمييز بين هاتين الحالتين- أي الحالة التي تمثل فيها المربعات الصغرى الاعتيادية ثاني أفضل خيار مقابل الحالة التي يكون فيها استخدامها غير صحيح تماماً. فمن الضروري في هذا الصدد إجراء المزيد من البحث لتحديد وفهم كيفية إجراء الدراسة التطبيقية في هذه الحالات المختلفة، لاسيما تلك الحالات التي لاينبغى استخدام المربعات الصغرى الاعتيادية فيها إطلاقاً.

ليس من أهداف هذا الفصل تطوير أساليب وطرق للتعامل مع أي مشكلة. فالمعالجة الناجحة لكافة هذه الحالات ستحتاج إلى العديد من المراجع المنهجية التي تتجاوز الإطار التمهيدي لهذا الكتاب! يهدف هذا الفصل فقط إلى تقديم حالات عامة قليلة ومصطلحاتها لتكون قادراً، على الأقل، على تنظيم حالات المشكلة عند بروزها وكذلك معرفة أين يمكنك البحث عن المزيد من الدراسة.

يمكن تقسيم المشكلات بصفة عامة إلى ثلاثة مجالات، سنقوم بمناقشة كل مجال على حدة أدناه:

(1) المشكلات التي تحدث عندما يكون للمتغير التابع أشكال معينة.

(2) المشكلات التي تحدث عندما يكون للأخطاء أشكال معينة.

(3) المشكلات التي تستدعي استخدام نماذج المعادلات المتعددة.

فإذا تجاوزت الأساليب الموضحة في الفصول السابقة وتعرفت على النماذج وأساليب التقدير الجديدة الموضحة في هذا الفصل، فسوف تدرك فوراً أنه سيكون من الصعب الاستمرار في استخدام برنامج إكسل. ولحسن الحظ، يتمتع العديد من برامج الاقتصاد القياسي المتخصصة المتوفرة في السوق (مثل:

Stata, LIMDEP, Gauss, E-views, TSP, MicroFit, PcGive, SHAZAM) بإمكانات وقدرات أفضل. فإذا كنت تخطط للقيام بقدر كبير من العمل التطبيقي في المستقبل، فيجب عليك تعلم كيفية استخدام برنامج واحد أو أكثر من هذه البرامج.

المشكلات التي تحدث عندما يكون للمتغير التابع أشكال معينة:

انظر نموذج الانحدار البسيط التالي:

$$Y = \alpha + \beta X + e.$$

ربما لاحظت، في الفصول السابقة، أننا قد ركزنا كثيراً على الحالات التي تعد فيها Y رقماً حقيقياً يأخذ أي قيمة (مثلاً: سعر مبيعات منزل، التغير النسبي في مساحة الغابات، نصيب الفرد من إجمالي الناتج المحلي). وقد تكون بلا شك قد واجهت، أو قد تواجه في المستقبل، بحالات يكون فيها للمتغير التابع شكل أكثر تقييداً، وهذا الشكل له انعكاسات على التقدير. وفيما يلي قائمة بالحالات الأكثر شيوعاً:

1) Y عبارة عن متغير صوري: لقد تم بحث هذه الحالة في نهاية الفصل السابع. لإعطاء فكرة موجزة، تبرز هذه الحالة عادةً عندما يكون المتغير التابع اختيارياً (مثلاً استخدام النقل العام أم لا، اختيار مهنة محددة أم لا). عندما يأخذ المتغير التابع هذا الشكل، فمن المحتمل أن يكون تقدير المربعات الصغرى الاعتيادية ملائماً، على الأقل، لوصف عام للمعلومات في البيانات، ولكن يوجد نماذج تقدير أفضل وتمثل نماذج Probit وLogit اثنين منها. لاحظ أن مصطلحات نماذج "المتغير التابع المحدود" (limited dependent variable) "الخيار النوعي" (qualitative choice)، "المتغير التابع المتقطع" (discrete dependent variable) و "المتغير التابع الصوري" (dummy dependent variable) في الكتب الأخرى تشير إلى هذه الحالة. وهذه هي عناوين الموضوعات التي يجب البحث عنها في أي مرجع علمي في مجال الاقتصاد القياسي إذا رغبت في تعلم المزيد عن هذه الأنواع من النماذج. بالإمكان أيضاً استخدام مثل هذه الأساليب في تناول الحالات التي يكون فيها Y عبارة عن نسبة (أي أن Y رقم يقع بين صفر وواحد).

2) قيمة Y مقيدة (محددة: هناك حدود للقيم التي يمكن أن تأخذها Y): تبرز هذه الحالات عندما تكون قيم Y أكبر أو أصغر من نقطة محددة. فعلى سبيل المثال، من الشائع جداً في مسوحات الدخل تسجيل دخل كل شخص في مسح استهلاك الأسرة عدا فئة السكان الأكثر ثراء والتي عادة ما يزيد دخلها عن مبلغ معين (100.000 جنيه مثلاً). لذلك، فإن الشخص الذي يبلغ دخله 20.000 جنيه سيسجل دخله على أساس 20.000 جنيه ولكن الذي يحصل على 200.000 جنيه سيسجل دخله على أساس 100.000 جنيه. ومثال ثان، افترض أن المتغير التابع هو مستوى الاستثمار المطلوب في شركة ولكن كل الذي تستطيع مشاهدته هو الاستثمار الفعلي بواسطة الشركة. فإذا

كان الاستثمار الفعلي لا يمكن أن يكون سالباً، فإن جميع الشركات بمستوى استثمار سالب سيتم تسجيلها على أن مستوى استثمارها صفر.

إذا كان المتغير التابع خاضعاً لقيود، فإن تقدير المربعات الصغرى الاعتيادية قد يكون مضللاً؛ أي أن المربعات الصغرى الاعتيادية سوف تكون "متحيزة" من منظور إحصائي[1]. وسوف تتزايد درجة التحيز مع نسبة المشاهدات الخاضعة للرقابة. فإذا كانت كمية قليلة من نقاط البيانات خاضعة للرقابة، فإن استخدام المربعات الصغرى الاعتيادية قد يكون مقبولاً. ولكن إذا كانت نسبة المشاهدات الخاضعة للقيود مرتفعة، فإنه ينبغي حتماً عدم استخدام المربعات الصغرى الاعتيادية. وفي مثل هذه الحالات، فإن الممارسة التطبيقية هي استخدام نماذج توبيت للتقدير (Tobit Estimator). يتم في مراجع الاقتصاد القياسي العلمية عادة معالجة مثل هذه الحالة تحت عنوان: "نماذج المتغيرات التابعة المحدودة".

3) Y عدد صحيح غير سالب: تبرز هذه الحالة عادة عندما يكون المتغير التابع هو عدد مرات وقوع الحدث. فعلى سبيل المثال، فإن Y قد يقيس عدد الوحدات المعطوبة في عملية إنتاج في أسبوع معين. وكبديل، فإن Y قد يقيس عدد براءات الاختراع بواسطة الشركة خلال سنة معينة. فإذا كان المتغير التابع بهذه الصيغة، فإن طريقة المربعات الصغرى الاعتيادية قد تكون ملائمة وإن كانت هناك أساليب أفضل. إن عبارة "نماذج بيانات معدودة أو محسوبة" (Count data models) هي التي ينبغي عليك البحث عنها في المراجع العلمية إذا كنت مهتماً بهذا النوع من دراسات الاقتصاد القياسي.

4) Y يقيس المدة "Duration": هذه الحالة كثيرة الشيوع في اقتصاديات العمل حيث يمثل المتغير التابع غالباً الوقت المستغرق في حالة معينة. فعلى سبيل المثال، قد يكون خبير اقتصاديات العمل مهتماً بتفسير سبب حصول بعض العاطلين على عمل أسرع من غيرهم. ففي هذه الحالة، فإن المتغير التابع هو

فترة البطالة لكل فرد (أي الوقت الذي يمضي قبل حصول الفرد على العمل). ثمة مثال آخر شائع الحدوث يتضمن تحليل الإجراءات الصناعية حيث يتمثل المتغير التابع في فترة الإضرابات. وفي مثل هذه الحالات، لا يكون تقدير المربعات الصغرى الاعتيادية للانحدار خاطئاً على نحو تلقائي. ومع ذلك، تم تطوير نماذج أخرى أفضل ربما ينبغي استخدامها. إن عبارة "نماذج المدة" (Duration models) هي التي يجب عليك البحث عنها في المراجع العلمية إذا كنت مهتماً بهذا النوع من دراسات الاقتصاد القياسي.

من المحتمل أن تكون الحالات الأربع المذكورة أعلاه هي الأكثر شيوعاً في هذه الفئة الخاصة من المشكلات الاقتصادية القياسية. إن استخدام طريقة المربعات الصغرى الاعتيادية يعد مضللاً فقط في حالة المتغير التابع الخاضع للقيود. ورغم ذلك، سيكون من الأفضل في مثل هذه الحالات قراءة المزيد حول الموضوع مع النظر في إمكانية استخدام برامج حاسب آلي غير إكسل.

المشكلات التي تحدث عندما يكون للأخطاء أشكال معينة:

لم نتناول حدود الخطأ في هذا الكتاب بأكثر من القول بأنها تقيس المسافة بين المشاهدة وخط الانحدار. هناك كم هائل من الدراسات التي توضح أي مُقدر هو الأفضل عند وضع الافتراضات بشأن الأخطاء. لن نناقش هنا النظرية الإحصائية الأساسية لهذه الأساليب بأي قدر من التفصيل. وبدلاً عن ذلك، سنقدم شرحاً كافياً حول أسباب حدوث هذه المشكلات وتوضيح أي الكلمات ينبغي عليك البحث عنها في مرجع منهجي في الاقتصاد القياسي إذا كنت ترغب في إجراء أبحاث مستقبلية في هذا المجال.

سنبدأ بالإشارة إلى أن المربعات الصغرى الاعتيادية هي المقدّر الأكثر استخداماً في نموذج الانحدار. إنها فعّالة للغاية في ظل العديد من الافتراضات

الإحصائية المختلفة. ورغم ذلك، فإنها تعد مثالية فقط إذا كانت كافة الأخطاء لها نفس الخصائص[2]. وقد تحصل، في حالات أخرى، على تقديرات دقيقة باستخدام مقدر "المربعات الصغرى المعممة" (Generalized least squares) أو (GLS). ولتعزيز فهم مقدر المربعات الصغرى المعممة (GLS)، سنبدأ بدراسة مشكلة اختلاف التباين (heteroskedasticity).

افترض أنك مهتم بتقدير انحدار النمو لعدد من الدول حيث يعتبر المتغير التابع "متوسط معدل نمو إجمالي الناتج المحلي" في كل دولة من أصل N دولة. تتضمن المتغيرات التفسيرية مستوى التعليم، الاستثمار، معدل الادخار،...الخ لكل دولة. افترض أن القائمة تتضمن العديد من الدول المتقدمة (الولايات المتحدة، المملكة المتحدة، ألمانيا، مثلاً) وكذلك العديد من الدول الأقل تقدماً (مثل السودان، أنجولا، هاييتي). فالدول المتقدمة لديها هيئات إحصاء (مؤسسات) حكومية جيدة التمويل تضطلع بمهمة جمع البيانات عن إجمالي الناتج المحلي، ومن ثم فإن بيانات إجمالي الناتج المحلي تكون دقيقة. وعلى العكس من ذلك، لا تتوافر في الدول الأقل تقدماً هيئات إحصاء جيدة التمويل لجمع البيانات ويكون جزء كبير من اقتصادياتها غير رسمية أو اقتصاديات فقيرة. نتيجة لذلك، تكون بيانات إجمالي الناتج المحلي الرسمية غير دقيقة. ما هي تبعات هذه الخصائص المتعلقة بجمع البيانات على تقدير المربعات الصغرى الاعتيادية؟

أولاً: إنها تعني أن الأخطاء، e، للدول الأقل تقدماً قد تكون أكبر بالمقارنة مع الدول المتقدمة. وهذه الظاهرة تعرف باسم "اختلاف التباين"[3].

ثانياً: قد نرغب في إعطاء وزن ترجيحي أكبر للدول المتقدمة نظراً لأن بياناتها قد تكون أكثر دقة من بيانات الدول الأقل تقدماً، وهذا هو عين ما يقوم به مقدر المربعات الصغرى المعممة. وفي الحقيقة، يمكن تفسير مقدر المربعات الصغرى المعممة بأنه مقدر المربعات الصغرى الاعتيادية باستخدام بيانات ذات

أوزان ترجيحية جديدة. فبدلاً عن استخدام Y و X في المعادلة لمقدر المربعات الصغرى الاعتيادية، فإن المربعات الصغرى المعممة تقوم باستحداث بيانات جديدة بصورة ضمنية (Y^*و X^* مثلاً) تعطي أوزاناً ترجيحية أكثر للدول المتقدمة وأوزاناً أقل للدول الأقل تقدماً.

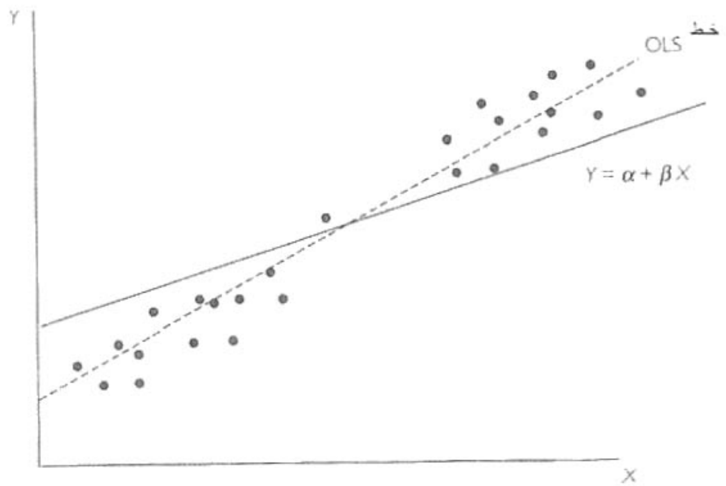
إذا كنت كثير البحث في الاقتصاد القياسي، فسوف تعرف بصورة دقيقة ماذا نعني بكلمة "إعادة الترجيح" (reweight). سيكون كافياً في هذه المرحلة ملاحظة أن المربعات الصغرى الاعتيادية لا تزال تمثل مقدراً جيداً[4]، ولكن المربعات الصغرى المعممة أفضل في حالة اختلاف التباين. بإمكانك إعطاء أوزان جديدة للبيانات واستخدام المربعات الصغرى المعممة في برنامج إكسل، ولكن ذلك يبدو غير ملائم مع معظم ملفات البيانات وسوف يكون من الأفضل استخدام برنامج حاسب آلي آخر.

بالإضافة إلى اختلاف التباين، هناك مشكلات أخرى ذات تبعات مماثلة على اختيار الأساليب في عملية التحليل (أي أن المربعات الصغرى الاعتيادية مناسبة ولكن المربعات الصغرى المعممة أفضل). وتحدث المشكلات الأكثر شيوعاً في هذا المجال عندما تكون الأخطاء مترابطة ذاتياً. لقد تم تعريف مفهوم الارتباط الذاتي في الفصل التاسع. ولكن الاختلاف هنا هو أن e_t وليس Y_t هو المترابط مع متباطئاته الذاتية. فإذا نجحت في اختبار طول المتباطئة الصحيح (أي p في نموذج AR(p) و p و q في نموذج ADL(p,q) فمن غير المحتمل أن يسبب الارتباط الذاتي للأخطاء أية مشكلة. ومع ذلك، ففي حالة أن يصبح مشكلة، فإن تقدير المربعات الصغرى المعممة سوف يحقق تقديرات أكثر دقة من تلك التي تحققها المربعات الصغرى الاعتيادية. ولا يستطيع برنامج إكسل، في هذه الحالة، تنفيذ المربعات الصغرى الاعتيادية بسهولة، ولكن العديد من برامج الحاسب الآلي الخاصة بالاقتصاد القياسي تكون قادرة على عمل ذلك بصورة مباشرة.

إن كافة الحالات المذكورة أعلاه هي حالات قد تكون فيها المربعات الصغرى الاعتيادية ملائمة وإن كانت المقدرات الأخرى أكثر ملاءمة للتطبيق. ورغم ذلك، هناك حالة واحدة مهمة حيث تعني مشكلات الأخطاء أن استخدام المربعات الصغرى الاعتيادية يكون مضللاً. وتحدث هذه الحالة عندما يكون الخطأ مرتبطاً بالمتغيرات التفسيرية. ولن تدرك كثيراً الأسباب التي تجعل هذه الحالة تسبب مشكلات ومتى قد تحدث. يوضح الشكل البياني رقم (12-1) حالة للارتباط الموجب بين الأخطاء والمتغيرات التفسيرية، وتم توضيح خط الانحدار الأصلي في الشكل البياني بخط متصل ($Y = \alpha + \beta X$). إن الارتباط الموجب بين X والأخطاء يعني أن قيم X المرتفعة مرتبطة مع الأخطاء (الموجبة) المرتفعة وأن قيم X المنخفضة مرتبطة مع الأخطاء (السالبة) المنخفضة. لذلك، سيكون الرسم البياني للمحورين XY من النوع الموضح في الشكل رقم (12-1) مع وقوع نقاط البيانات أسفل خط الانحدار بالنسبة لقيم X المنخفضة وفوق الخط لقيم X المرتفعة. بعبارة أخرى، تم رسم الشكل رقم (12-1) على النحو الذي يجعل كافة الأخطاء سالبة لقيم X المنخفضة وموجبة لكافة قيم X المرتفعة. وتأخذ المربعات الصغرى الاعتيادية خط ملاءمة عبر نقاط البيانات ينتهي بالخط المسمى OLS في الشكل رقم (12-1). ونظراً لأن هذا الخط المتقطع يأخذ ميلاً مختلفاً وقاطعاً مختلفاً عن خط الانحدار الحقيقي، فمن الواضح أن المربعات الصغرى الاعتيادية غير ملائمة في هذه الحالة.

يعد نموذج المعادلات الآنية (simultaneous equations) الحالة الأكثر شيوعاً حيث تكون المتغيرات التفسيرية والأخطاء مترابطة، وسيتم بحث هذا الموضوع في الجزء القادم. ولكن يجب عليك في هذه المرحلة فقط ملاحظة أنه إذا كانت الأخطاء مرتبطة مع المتغيرات التفسيرية، فلا ينبغي عليك استخدام

المربعات الصغرى الاعتيادية. بدلاً عن ذلك، يجب عليك تعلم نماذج المتغيرات المساعدة (instrumental variables) وكيفية استخدامها.



شكل رقم (12-1): المربعات الصغرى الاعتيادية عندما يكون الخطأ مرتبطاً مع قيمة **X**.

المشكلات التي تستدعي استخدام نماذج المعادلات المتعددة:

لقد ركزنا، في كافة أجزاء هذا الكتاب، على نموذج الانحدار ذي المعادلة الواحدة. ولكن من الشائع عملياً، التعامل مع عدة متغيرات تابعة ومن ثم عدة نماذج انحدار (تسمى أحياناً "نظم المعادلات" systems of equations). وفيما يلي بعض الأمثلة لتوضيح متى تبرز هذه التطبيقات.

1- تخيل أنك قد جمعت بيانات مقطعية عن حجم الإنتاج في عدد من الشركات. بالتحديد سوف تتضمن بياناتك عدد العمال، حجم رأس المال، الطاقة والمدخلات من المواد والسعر لكل مدخل. سيكون اهتمامك منصباً على تفسير خيارات المدخلات المتاحة للشركة. وفي هذه الحالة، لديك أربعة متغيرات تابعة مختلفة (أي العمل، رأس المال، الطاقة، المواد) تعتمد جميعها

على أسعار المدخلات. سيتعين عليك تقدير معادلة انحدار منفصلة أو مستقلة لكل متغير تابع ومن ثم ستحصل على أربع معادلات مختلفة.

2- تخيل أن لديك بيانات سلسلة زمنية عن الاستهلاك مفصلة حسب القطاعات (مثلاً، استهلاك الأغذية، النقل، السكن، الملبس، السلع المعمرة،... الخ). إن اهتمامك هو معرفة كيفية اعتماد هذه المكونات المختلفة للاستهلاك على وضع الاقتصاد. وفي هذه الحالة، ستكون لديك العديد من المتغيرات التابعة (مثلاً، استهلاك الأغذية، النقل، المسكن، ... الخ) وسوف تستخدم متغيرات اقتصادية كلية مثل إجمالي الناتج المحلي، أسعار الفائدة، ... الخ كمتغيرات تفسيرية. إن كل متغير تابع يعني وجود معادلة انحدار مختلفة.

3- في الدراسة المالية، قد تكون مهتماً بتفسير عائدات أسهم العديد من الشركات. وفي هذه الحالة، ستكون لديك عدة متغيرات تابعة مختلفة (أي عائد سهم كل شركة) معتمدة على متغيرات تفسيرية مثل سعر الفائدة، ... الخ.

4- إن نموذج متجه الانحدار الذاتي (VAR)، الذي تم بحثه في الفصل الحادي عشر، هو نموذج معادلات متعددة. ففي هذا النموذج، يعتمد Y على متباطئاته الذاتية ومتباطئات المتغير الآخر، X. ولهذا النموذج أيضاً معادلة أخرى يكون فيها X المتغير التابع ويعتمد الأخير على متباطئاته الذاتية ومتباطئات Y.

5- يعتبر نموذج (IS-LM) معروفاً جداً لدارسي الاقتصاد الكلي. لاحظ أن هذا النموذج قائم على معادلتين هما: LM و IS.

6- يعد نموذج العرض والطلب أيضاً نموذجاً شائعاً في علم الاقتصاد. ولهذا النموذج معادلتان: إحداهما دالة منحنى العرض والأخرى دالة منحنى الطلب.

إن النقطة الأساسية لهذا الجزء هي أن المثالين الأخيرين فقط هما المسببان لمشكلات رئيسة لتقدير المربعات الصغرى الاعتيادية.

ولفهم المشكلات التي تبرز في نظم المعادلات المتعددة، يجب علينا إدخال بعض المفاهيم التي ربما تكون على دراية بها لدراستك السابقة للاقتصاد. فالمتغير يعد متغيراً داخلياً (Endogenous Variable) إذا تم تحديده ضمن النموذج قيد الدراسة، في حين يكون متغيراً خارجياً (Exogenous Variable) إذا لم يتم تحديده ضمن النموذج.

ترتبط هذه المفاهيم ارتباطاً وثيقا بقضايا السببية التي نوقشت في الفصل الرابع. تذكر أننا أكدنا أن تفسير الانحدار يكون أسهل إذا كان المتغير التفسيري سبباً في حدوث المتغير التابع (وليس العكس). بعبارة أخرى، يفترض نموذج الانحدار أن Y يتحدد بما يحدث لـ X. إننا لم نحدد الكيفية التي حدث بها X. في هذه الحالة، يعد المتغير التابع Y بمثابة المتغير الداخلي مع افتراض المتغير التفسيري X بمثابة متغير خارجي. وما دامت متغيراتك التفسيرية خارجية، فإن استخدام تقدير المربعات الصغرى الاعتيادية يعد جيداً حتى في حالة وجود معادلات متعددة. ولكن إذا كانت المتغيرات التفسيرية داخلية، فلا ينبغي استخدام تقدير المربعات الصغرى الاعتيادية.

وفيما يلي بعض الأمثلة الضرورية لتوضيح المزيد حول هذه الأفكار:

1- في الفصلين السادس والسابع، أجرينا انحدار أسعار المنازل على مجموعة متنوعة من مواصفات المنزل. هنا تعتمد الأسعار، Y، على مواصفات المنزل، X (مثل، المنزل الذي يضم عدداً أكثر من غرف النوم يكون سعره أكبر). ولكن مواصفات المنزل لا تعتمد على سعره (مثلا، إذا انهارت سوق العقار وانخفض سعر المنزل، فإن ذلك لن يجعل للمنزل عدداً أقل من غرف

النوم أو الحمامات). إن X يسبب Y ولكن Y لا يسبب X؛ (X خارجي و Y داخلي).

2- تختار الشركات مستويات مدخلاتها وفقاً لأسعار هذه المدخلات (مثلاً: إذا كانت الأجور منخفضة بالمقارنة مع سعر شراء آلة جديدة، فإن الشركات سوف تميل إلى توظيف عمّال بدلاً عن شراء آلات جديدة). فأسعار الأخيرة تحدد أو تسبب خيار المدخل. ولكن خيار المدخل للشركة لا يؤثر في سعر المدخلات. فعلى سبيل المثال، وفي الأسواق التنافسية على أقل تقدير، إذا وظفت الشركة المزيد من العمّال، فإن هذا الإجراء لا يسبب ارتفاع الأجور. لذلك فإن المدخلات في نموذج مصمم لتفسير خيار المدخل، ستكون داخلية (يحددها النموذج) ولكن أسعار المدخلات ستكون خارجية.

3- إذا قمت بحل نموذج IS-LM، فسوف تحصل على إجابات نموذجية للدخل وسعر الفائدة. بمعنى أن الدخل وسعر الفائدة يتم تحديدهما (أو حلهما) بواسطة النموذج. وعندما يسعى خبراء الاقتصاد القياسي إلى تقدير نماذج IS-LM، فإنهم يستخدمون هذين المتغيرين (ضمن متغيرات أخرى). لاحظ أن كليهما داخلي (يحددهما النموذج).

4- يحدث التوازن بين السعر والكمية، في نموذج العرض والطلب، عند نقطة تقاطع منحنى العرض مع منحنى الطلب. يتم تحديد سعر وكمية السلع المعروضة والمطلوبة في السوق بواسطة النموذج. ومن ثم فإن كلاً من السعر والكمية داخليان.

أما فيما يتعلق بباقي هذا الجزء فسنفترض أن Y متغير داخلي و X متغير خارجي. فإذا كان لدينا أكثر من واحد من هذين المتغيرين، فسوف تستخدم

الرموز $Y_1,...,Y_M$ و $X_1,...,X_K$ للإشارة إلى عدد M متغير داخلي وK متغير خارجي. وفيما يلي تصنيف إجمالي للحالات المحتملة:

1- لقد تم بحث نموذج الانحدار $Y = \alpha + \beta_1 X_1 + ... + \beta_K X_K + e$ بالتفصيل في هذا الكتاب. وبالإمكان إجراء تقدير المربعات الصغرى الاعتيادية بسهولة[5].

2- إذا كانت لديك منظومة معادلات بالصيغة التالية:

$$Y_1 = \alpha_1 + \beta_{11} X_1 + ... + \beta_{1K} X_K + e_1$$

$$Y_2 = \alpha_2 + \beta_{21} X_1 + ... + \beta_{2K} X_K + e_2$$

....

....

$$Y_M = \alpha_M + \beta_{M1} X_1 + ... + \beta_{MK} X_K + e_M$$

قم بإجراء المربعات الصغرى الاعتيادية لمعادلة واحدة في كل مرة. لاحظ أن هذا النموذج يفترض أن كافة المعادلات لها المتغيرات التفسيرية ذاتها تماماً (مثلا: تعتمد كمية كل مدخل على أسعار كافة المدخلات). فإذا كانت المتغيرات التفسيرية للمعادلات مختلفة (مثلا: تعتمد كمية كل مدخل فقط على سعره)، فإن هناك أسلوب تقدير أفضل من المربعات الصغرى الاعتيادية وهو أسلوب معادلات الانحدار غير المترابطة ظاهرياً Seemingly Unrelated regression equations (SURE).

ولكن لاحظ أن المربعات الصغرى الاعتيادية تمثل ثاني أفضل أسلوب تقدير ملائم.

3- إذا كنت تتعامل مع النموذج التالي:

$$Y_1 = \alpha + \beta Y_2 + e$$

أو على نحو أكثر عموماً:

$$Y_1 = \alpha_1 + \gamma_{12}Y_2 + ... + \gamma_{1M}Y_M + \beta_{11}X_1 + ... + \beta_{1K}X_K + e_1$$

$$Y_2 = \alpha_2 + \gamma_{21}Y_1 + ... + \gamma_{2M}Y_M + \beta_{21}X_1 + ... + \beta_{2K}X_K + e_2$$

....

....

$$Y_M = \alpha_M + \gamma_{M1}Y_1 + ... + \gamma_{M,M-1}Y_{M-1} + \beta_{M1}X_1 + ... + \beta_{MK}X_K + e_M$$

فإن استخدام المربعات الصغرى الاعتيادية قد يعطي نتائج مضللة ومن ثم يجب تجنب ذلك.

خلاصة القول أنه إذا كانت المتغيرات التفسيرية خارجية، فسوف يكون استخدام المربعات الصغرى الاعتيادية (حتى في نماذج المعادلات المتعددة) مقبولاً. أما إذا كانت المتغيرات التفسيرية داخلية، فإن المربعات الصغرى الاعتيادية لا تعد أسلوباً مناسباً للتقدير (حتى في نموذج أحادي المعادلة). يعرف النموذج الأخير أعلاه بنموذج المعادلات الآنية (simultaneous equations model) الذي حظي بقدر هائل من الاهتمام في دراسات الاقتصاد القياسي. ولكن بحث هذا النموذج بالتفصيل يعد خارج نطاق هذا الكتاب. ورغم ذلك، فمن الملائم التوضيح بإيجاز أسباب حدوث مشكلة مع المربعات الصغرى الاعتيادية من خلال دراسة المثال التالي:

انظر أبسط نسخة من نموذج العرض والطلب في علم الاقتصاد. يتم تحديد منحنى الطلب بالمعادلة التالية:

$$Q^D = \alpha_D + \beta_D P$$

وهذه المعادلة توضح أن الكمية المطلوبة من سلعة، Q^D، تعتمد على سعرها، P. وتوضح منحنى العرض كيفية اعتماد الكمية المعروضة بواسطة الشركات، Q^S، أيضا على السعر:

$$Q^S = \alpha_S + \beta_S P$$

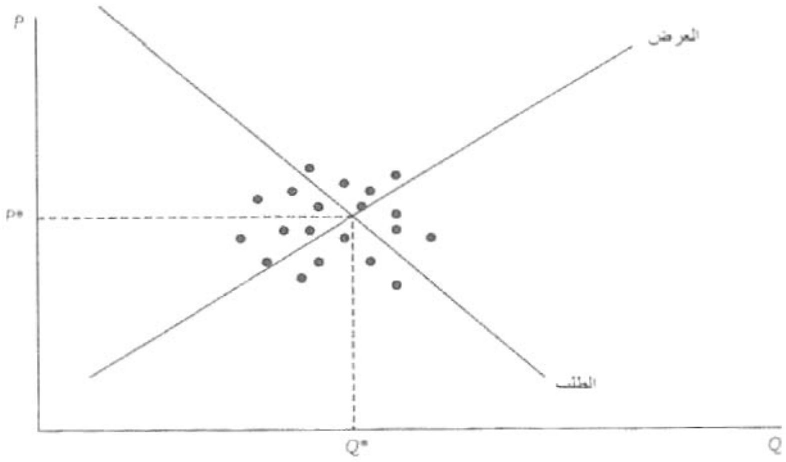
يوضح الخطان المتصلان في الشكل رقم (12-2) منحنى العرض ومنحنى الطلب. وتوضح نقطة تقاطع الخطين سعر وكمية التوازن P^* و Q^*. بعبارة أخرى، يتم تحديد السعر والكمية في النموذج، وهما داخليان.

ماذا سيحدث إذا تم الحصول على السعر والكمية (مثلاً: من السوق لمنتج معين كل أسبوع ولعدة أسابيع) وتم إجراء انحدار الكمية على السعر؟

كنا سنحصل على تقديرات المربعات الصغرى الاعتيادية للقاطع والميل، $\hat{\alpha}$ و $\hat{\beta}$ مثلاً. ولكن ماذا تمثل تقديرات $\hat{\alpha}$ و $\hat{\beta}$؟ ربما تواجه مشكلة في استخدام المربعات الصغرى الاعتيادية: إننا لانعرف ما إذا كانت $\hat{\alpha}$ و $\hat{\beta}$ هي تقديرات α_D و β_D (أي منحنى الطلب) أم α_S و β_S (أي منحنى العرض). ومن منظور لغة الاقتصاد القياسي، يعد هذا مثالاً لمشكلة تعريف النموذج (identification problem). ومن الناحية العملية، فإن المربعات الصغرى الاعتيادية قد لاتقدر منحنى العرض ولا منحنى الطلب.

لمزيد من التوضيح عن سبب فشل المربعات الصغرى الاعتيادية، سنلقي نظرة على الشكل رقم (12-2). يمثل P^* و Q^* سعر وكمية التوازن. افترض أننا نشاهد السعر والكمية لهذه السلعة في سوق ما لعدة مرات (مثلاً: أسبوعياً

لمدة سنة). ففي الواقع، فإنه من المحتمل أن لانكون في وضع التوازن الدقيق وأن بعض الأخطاء الصغيرة سوف تحدث. وهذا يعني أن السعر والكمية الفعلية المشاهدة كل أسبوع لن تكون P^* و Q^* تماماً كل مرة. ومن ثم فإن نقاط البيانات المشاهدة قد تكون متمركزة حول نقطة التوازن على النحو الظاهر في الشكل رقم (2-12). تخيل محاولة، كما في المربعات الصغرى الاعتيادية، رسم خط مستقيم يمر عبر هذه النقاط. من الواضح أن ذلك الخط لن تكون له علاقة، بالضرورة، بمنحنى العرض أو منحنى الطلب.



شكل رقم (2-12): البيانات المشاهدة في نموذج العرض والطلب

يمكننا على نحو أكثر منهجية توضيح أنه عندما يكون أي من المتغيرات التفسيرية، أو جميعها، داخلياً، فإن خطأ الانحدار سوف يكون مرتبطاً مع المتغيرات التفسيرية وأن استخدام المربعات الصغرى الاعتيادية سوف يكون خاطئاً (انظر الشرح الخاص بالشكل رقم 12-1 أعلاه). فإذا واجهت مثل هذه الحالة فيجب عليك إجراء المزيد من الدراسة واكتساب المزيد من المعرفة عن نماذج المعادلات الآنية وتقدير المتغيرات المساعدة.

ملخص الفصل:

على الرغم من أن تقدير المربعات الصغرى الاعتيادية يمثل أداة فعالة يمكن استخدامها لأشكال متنوعة من البيانات، إلا أنه ليس مثالياً في كل الحالات. فهناك بعض الحالات التي لايعد فيها استخدام المربعات الصغرى الاعتيادية الخيار الأفضل، وإن كان ملائماً. كما أن هناك حالات يعد فيها استخدام المربعات الصغرى الاعتيادية مضللاً للغاية. وفيما يلي قائمة موجزة لهذين النوعين من الحالات.

الحالات التي لا يعد فيها استخدام المربعات الصغرى الاعتيادية الخيار الأفضل:

1) المتغير التابع عبارة عن متغير صوري، مدة زمنية أو معدود (أي عدد صحيح).

2) تتسم الأخطاء باختلاف التباين أو الارتباط الذاتي (جرى شرح هذين المفهومين بالتفصيل في الفصل).

3) للبيانات متغيرات تابعة متعددة ومن ثم معادلات متعددة، ولكن كافة المتغيرات التفسيرية خارجية.

الحالات التي يجب فيها تجنب المربعات الصغرى لكونها مضللة:

1- قيمة المتغير التابع مقيدة (محددة).

2- الأخطاء مرتبطة مع المتغيرات التفسيرية.

3- يكون واحد، أو أكثر، من المتغيرات التفسيرية داخلياً.

4- للبيانات متغيرات تابعة متعددة ومن ثم معادلات متعددة، ولكن بعض المتغيرات التفسيرية داخلية.

ملاحظات ختامية:

1- التحيز (Bias) عبارة عن مصطلح إحصائي لن يتم تعريفه بصورة منهجية هنا. ولكن بصفة عامة، إذا استخدمت مقدراً غير متحيز في العديد من التطبيقات فإن تقديرك قد يكون مرتفعاً أو منخفضاً في أي تطبيق فردي ولكنه في المتوسط سيكون صحيحاً. وعلى النقيض من ذلك، ستكون المقدرات المتحيزة، في المتوسط، خاطئة؛ لذا يجب تجنبها.

2- المربعات الصغرى الاعتيادية، من المنظور الإحصائي، هي الأفضل من جميع المقدرات (في فئة محددة) إذا كانت الأخطاء مستقلة عن بعضها وكانت جميعها مستنتجة من التوزيع ذاته. فإذا كان هذا التوزيع طبيعياً، بالإمكان تقديم دليل قوي على أن تقدير المربعات الصغرى الاعتيادية هو الأمثل.

3- يحدث اختلاف التباين إذا كان الانحراف المعياري للأخطاء مختلفاً عبر المشاهدات المختلفة.

4- المربعات الصغرى الاعتيادية والمربعات الصغرى المعممة، من منظور إحصائي، تتكون من مقدرين غير متحيزين ولكن المربعات الصغرى المعممة أكثر كفاءة من المربعات الصغرى الاعتيادية.

5- بالطبع، إذا كانت مشكلات الخطأ أو المتغير التابع، التي تم بحثها في هذا الفصل، قائمة فإنه يجب تعديل المربعات الصغرى الاعتيادية وفقاً للنقاش السابق. وهذا يصح بالنسبة للحالات التالية أيضاً.

الملاحق

ملحق رقم (أ)

إعداد مشروع بحث قياسي (تطبيقي)

(Writing an Empirical Project)

يقدم هذا الملحق إرشادات عامة لكيفية إعداد مشروع بحث قياسي (تطبيقي). سيتبع هذه الإرشادات موضوعات لمشروعين (بما في ذلك البيانات المطلوبة لتنفيذهما) بهدف تحقيق فهم أعمق للأساليب التي تم وصفها في هذا الكتاب.

وصف مشروع بحث قياسي نموذجي:

يقوم اقتصاديو اليوم بإجراء البحوث في مجالات متنوعة. إن طلاب المرحلة الجامعية والدراسات العليا، والاقتصاديين الأكاديميين، وراسمي السياسات في المؤسسات الحكومية والبنوك المركزية، والاقتصاديين المهنيين العاملين في المصارف الخاصة أو المؤسسات الخاصة، قد يكونون جميعاً في حاجة إلى إعداد التقارير والبحوث والدراسات التي تتضمن تحليل بيانات اقتصادية.

واستناداً إلى نوع الموضوع والقراء المستهدفين، يختلف شكل هذه التقارير اختلافاً كبيراً بحيث لاتوجد صيغة واحدة ثابتة للتقرير القياسي. وسنقدم أدناه عناصر مشتركة للتقارير والدراسات الاقتصادية لتكون مرشداً للدراسات القياسية المستقبلية. ورغم ذلك، ينبغي ملاحظة أنه في سياق مشروعاتك الجامعية أو مجال عملك، ليس بالضرورة إدراج كافة هذه العناصر في تقريرك.

1- **مقدمة:** تبدأ معظم التقارير والبحوث والدراسات بمقدمة تتضمن وصفاً موجزاً للمشكلة قيد الدراسة وتلخص النتائج القياسية الرئيسية. يجب كتابة المقدمة بلغة سهلة وغير فنية مع أقل قدر ممكن من استخدام اللغة الإحصائية

والاقتصادية. فالقارئ غير المتخصص في المجال المعني، يجب أن يكون قادراً على قراءة وفهم المشكلات العامة ونتائج التقرير أو الدراسة.

2- **استعراض الدراسات السابقة:** يجب في هذا الجزء تلخيص البحوث والدراسات ذات العلاقة التي قام بها الآخرون وذلك من خلال حصر الدراسات السابقة والنتائج وثيقة الصلة بتقريرك وتقديم وصف موجز عنها.

3- **النظرية الاقتصادية:** إذا كانت دراستك ذات طابع أكاديمي وتتضمن نموذجاً نظرياً منهجياً، فغالباً ما يتم وصف ذلك في هذا الجزء. أما عن الدراسات المتعلقة بالسياسات، فلن تكون في حاجة لإدراج مثل هذا النموذج الرياضي المنهجي ولكن هذا الجزء يتيح لك وصف المشكلات الاقتصادية أو المؤسسية لعملك بتفصيل أكثر. وهذا الجزء قد يكون أكثر تخصصاً (فنياً) من الأجزاء السابقة وسوف يتضمن من الناحية النموذجية قدراً من المصطلحات الاقتصادية والرياضية. بإيجاز، يمكنك توجيه هذا الجزء حصرياً للقراء المتخصصين في هذا المجال.

4- **البيانات:** يجب عليك في هذا الجزء وصف البيانات، بما في ذلك شرح كامل لمصادر هذه البيانات.

5- **النموذج المطلوب تقديره:** يجب عليك في هذا الجزء تحديد كيفية استخدام البيانات لاختبار النظرية الاقتصادية الموضحة في الجزء رقم (3). إن الشكل الفعلي لهذا الجزء قد يختلف كثيراً حسب نوع الموضوع والقراء المستهدفين. فعلى سبيل المثال قد ترغب في إثبات أن انحداراً معيناً مهم للدراسة، أو أن متغيراً معيناً سيكون متغيراً تابعاً وأن متغيرات أخرى ستكون تفسيرية. وبالمثل في حالة سلسلة زمنية لمتغير اقتصادي كلي قد ترغب في إثبات أن النظرية الاقتصادية تؤكد ضمناً أن متغيراتك يجب أن تكون متكاملة

ومن ثم يجب إجراء اختبار تكامل. بإيجاز هذا هو الجزء الذي يجب عليك فيه إعطاء مبررات للأساليب المستخدمة في الجزء التالي.

6- **النتائج القياسية:** يمثل هذا الجزء عادة الجزء الأساسي لأي تقرير أو دراسة. ففي هذه المرحلة، يتعين عليك وصف النتائج القياسية التي توصلت إليها وبحث كيفية ربطها بالمشكلة (المشكلات) الاقتصادية قيد الدراسة. يجب أن يتضمن هذا الجزء المعلومات الاقتصادية والإحصائية. ونعني بالمعلومات "الاقتصادية"، مثلاً، تقديرات المعامل أو وجود تكامل بين متغيرين وماذا تعني هذه النتائج بالنسبة للنظرية الاقتصادية. وعلى العكس من ذلك، قد تتضمن المعلومات "الإحصائية": نتائج من اختبارات الفروض توضح مدى معنوية تقديرات المعامل، مبررات اختيار قيمة المتباطئة، تفسير استبعاد المتغيرات التفسيرية غير المعنوية، بحث ملاءمة النموذج (مثل R^2 أو القيم الشاذة)،... الخ. وبالإمكان عرض الكثير من هذه المعلومات في شكل رسوم بيانية. ويعد أمراً عادياً أن تبدأ الدراسات بأشكال بيانية بسيطة (مثل رسم بياني لسلسلة زمنية للبيانات) وإتباع ذلك بجدول إحصاءات وصفية (مثل المتوسط، الانحراف المعياري، والحد الأقصى/الأدنى لكل متغير، ومصفوفة ارتباط). وبالإمكان إضافة جدول آخر للنتائج من تحليل إحصائي أكثر منهجية مثل تقديرات معاملات المربعات الصغرى الاعتيادية وإحصائيات t (أو قيم P)، R_s^2 وإحصائية F لاختبار معنوية الانحدار ككل.

7- **خاتمة:** يجب في هذا الجزء تقديم ملخص موجز للمشكلات المعالجة في الدراسة بصفة عامة وأهم النتائج القياسية بشكل خاص وكذلك التوصيات.

الاعتبارات العامة:

يتضمن ما يلي وصفاً لعدد محدود من المشكلات التي ينبغي عليك أخذها في الاعتبار أثناء تنفيذ دراسة قياسية. ويتم التركيز خاصة على بحث مقومات دراسة تطبيقية علمية جيدة وكيفية عرض نتائجها.

إن أول ما ينبغي تأكيده هو أنه ليس هناك نتائج قياسية صحيحة أو خاطئة. النتائج القياسية تكون كما هي؛ لذا لا ينبغي الاستياء إذا لم تحصل على النتائج التي كنت تأملها. ففي العالم المثالي، يحصل الباحث على نظرية جديدة ثم يجري العمل التطبيقي الذي يدعم هذه النظرية الجديدة بطريقة ذات معنوية إحصائية. أما في الواقع العملي فنادراً ما يحصل هذا الوضع المثالي.

إن المتغيرات التفسيرية التي تتوقع أن تكون ذات معنوية إحصائية غالباً لا تكون كذلك في عالم الواقع. كما أن المعاملات التي تتوقع أن تكون موجبة قد يتضح أنها سالبة. ويتم الحصول على هذه النتائج باستمرار حتى في أكثر الدراسات تعقيداً؛ لذا الا ينبغي أن تثبط عزيمتك. على العكس، يجب أن تكون دائماً ذا عقلية منفتحة. فالنتيجة التي تقول أن النظرية ليست صحيحة يعد من الناحية العلمية صحيحة تماماً كالنتيجة التي تقول أن النظرية صحيحة.

بالإضافة إلى ذلك، تكون النتائج القياسية غالباً غير واضحة أو مربكة. فعلى سبيل المثال، قد يشير اختبار إحصائي إلى شيء ما، في حين يشير آخر إلى العكس. وبالمثل، فإن المتغير التفسيري الذي يعد معنوياً في انحدار ما قد يكون غير معنوي في انحدار آخر. ليس بإمكانك عمل شيء في هذا الصدد غير نقل نتائجك بأمانة وصدق ومحاولة (إذا أمكن) فهم أسباب حدوث مثل هذا التضارب أو عدم الوضوح.

من النادر جداً أن يقوم الاقتصاديون بتلفيق (عرض نتائج غير حقيقية) نتائجهم بالكامل. ورغم ذلك، فإنهم في الغالب يلجأون إلى القيام ببعض التصرفات غير النزيهة لإظهار أن النتائج جاءت كما الاستنتاجات الاقتصادية المتوقعة. فعلى سبيل المثال، من المعتاد قيام الباحثين بإجراء عدد كبير من الانحدارات مع العديد من المتغيرات التفسيرية المختلفة. عموماً، يعد هذا أمراً حكيماً، ودليلاً على أن الباحثين يدرسون ويعالجون خصائص البيانات بالتفصيل ومن زوايا مختلفة. ورغم ذلك، إذا عرض الباحثون فقط الانحدار الذي يدعم نظرية بعينها دون غيره من الانحدارات التى لاتدعمها، فإن ذلك يعني تضليل القراء عن قصد. حاول دائماً تجنب مثل هذه النزعة التى تهدف إلى سوء عرض النتائج.

وفيما يختص بكيفية عرض النتائج، ينبغي التأكيد على أهمية الوضوح والإيجاز. فمما لاشك فيه أن كلاً من أساتذة الجامعات، موظفي الخدمة المدنية، راسمي السياسات، أصحاب العمل دائماً منشغلون ولا يرغبون في إضاعة الكثير من الوقت في قراءة تقارير مطولة وغير مرتبة.

تعد الانتقائية إحدى المهارات الأساسية التى ينبغي أن يتمتع بها كاتبو التقارير الجيدة. فعلى سبيل المثال، قد تكون لديك نتائج معاملات مختلفة وإحصائيات اختبارات عديدة من عمليات إجراء الانحدارات المختلفة. إن جزءاً مهماً من أي تقرير هو تحديد أي معلومة هي المهمة وأيها غير مهمة بالنسبة للقراء المستهدفين. اختر فقط المعلومات الأكثر أهمية لإدراجها في التقرير وحاول دائماً عرض النتائج التي حصلت عليها بأمانة ووضوح.

موضوعات المشروع:

فيما يلي نعرض موضوعي مشروعين قد ترغب في تنفيذهما:

المشروع الأول: لغز التقييم المنخفض لأسعار الأسهم:

خلفية:

يهتم المستثمرون والاقتصاديون الماليون بفهم كيفية تقييم أسهم شركة ما. ينبغي أن تعكس قيمة أسهم الشركة، في الغالب، تطلعات المستثمرين فيما يتعلق بالأرباح المستقبلية للشركة. ولكن البيانات المتعلقة بالربحية المستقبلية المتوقعة غير متوافرة. بدلاً عن ذلك، يجب على الدراسات المالية التطبيقية استخدام مقاييس مثل الدخل الحالي، المبيعات، الموجودات والديون الخاصة بالشركة كمتغيرات تفسيرية.

بالإضافة إلى التساؤل العام بشأن كيفية تقييم أسواق الأسهم للشركات، هناك سؤال آخر حظي بقدر كبير من اهتمام الاقتصاديين الماليين في الآونة الأخيرة. لتوضيح وفهم هذه المشكلة، لاحظ أن معظم الأسهم المتداولة في سوق الأسهم هي أسهم قديمة للشركات القائمة. ولكن العديد من الشركات القديمة سوف تصدر أسهماً جديدة بالإضافة لتلك المتداولة- أى ما يعرف بـ "الإصدار الثاني (الإضافي) للأسهم" أو (SEO). علاوة على ذلك، قد تقرر بعض الشركات، التى لم تطرح أسهمها للتداول من قبل، إصدار مثل هذه الأسهم الآن (مثلا: قد تقرر شركة برامج حاسب آلي يملكها أحد الأفراد تحويلها إلى شركة عامة وبيع أسهمها للحصول على الأموال اللازمة للاستثمار المستقبلي أو التوسعة). تسمى مثل هذه الأسهم بـ "الإصدار العام الأولي" أو (IPO). ويرى بعض الباحثين استناداً إلى بيانات مالية، أن الإصدار العام الأولي يكون أقل قيمة بالمقارنة مع الإصدار الثاني (الإضافي) للأسهم (على الرغم من أن آخر دراسة في هذا الصدد قد أشارت إلى العكس).

في هذا المشروع، مطلوب منك إجراء دراسة قياسية للإجابة عن هذه الأسئلة باستخدام ملف البيانات التالي.

البيانات:

يحتوي ملف إكسل EQUITY.XLS بيانات لعدد من الشركات (N=309) شركة قامت بإصدار أسهم جديدة عام 1996. بعض هذه الأسهم ما هي إلا إصدار ثان للأسهم (SEO) والبعض الآخر إصدار أولي (IPO). تم توفير البيانات عن المتغيرات التالية، مع ملاحظة أن كافة المتغيرات عدا SEO مقاسة بملايين الدولارات:

- VALUE= القيمة الإجمالية لجميع الأسهم (الجديدة والقديمة) بعد إصدار الشركة للأسهم الجديدة. يتم حساب ذلك بضرب سعر السهم في عدد الأسهم.
- DEBT= قيمة الدين طويل الأجل على الشركة.
- SALES= إجمالي مبيعات الشركة.
- INCOME= صافي دخل الشركة.
- ASSETS= القيمة الدفترية لأصول الشركة (أي تقييم المحاسب لقيمة أصول الشركة).
- SEO= متغير صوري يساوي واحداً إذا كانت الأسهم المصدرة عبارة عن SEO ويساوي صفراً إذا كانت عبارة عن IPO.

المشروع الثاني: سلوك تحديد الأجور

خلفية:

يتيح لك هذا المشروع إمكانية اختبار سلوك تحديد الأجور باستخدام بيانات السلاسل الزمنية. إن المشكلة العامة محل الاهتمام في مثل هذه التحليلات هي كيف تتأثر الأجور بالعوامل الاقتصادية الكلية مثل مستوى الأسعار، إجمالي الناتج المحلي والمتغيرات التي تعكس معدل البطالة والتوظيف. يجب أن يتضمن

التحليل القياسي لمثل هذه البيانات بحث مشكلات مثل جذور الوحدة والتكامل المشترك.

البيانات:

يحتوي ملف إكسل WAGE.XLS على بيانات سنوية عن المملكة المتحدة للفترة 1855-1987. تم حساب اللوغاريثمات لجميع المتغيرات. ويتضمن الملف البيانات عن المتغيرات التالية:

- W= لوغاريثم الأجور الاسمية.
- P= لوغاريثم الرقم القياسي لأسعار السلع الاستهلاكية.
- GDP= لوغاريثم إجمالي الناتج المحلي الحقيقي.
- E= لوغاريثم إجمالي العمالة الفعلية.
- L= لوغاريثم إجمالي القوى العاملة المتاحة.

معلومات إضافية:

بالإضافة إلى المشكلة العامة لسلوك تحديد الأجور، يتركز الاهتمام الاقتصادي على بعض وظائف المتغيرات الواردة هنا. إذا كنت تتذكر خواص الصيغة اللوغاريثمية مثل:

$$\ln(\frac{A}{B}) = lN(A) - \ln(B)$$

و

$$\ln(1 + A) \approx A$$

بإمكانك اشتقاق العلاقات التالية:

- لوغاريثم الأجور الحقيقية ≈ W-P
- لوغاريثم إنتاجية العامل ≈ GDP-E
- لوغاريثم معدل البطالة ≈ L-E
- لوغاريثم نسبة الأجور من إجمالي الناتج المحلي = W-P-GDP+E

تتمثل إحدى المشكلات التي قد تكون مهتماً بدراستها، في تحديد ما إذا كانت العلاقات الموضحة أعلاه متكاملة ذاتياً. لقد أوضحنا في الفصل الحادي عشر تقدير انحدار متكامل ذاتياً باستخدام أساليب المربعات الصغرى الاعتيادية- وقد ترغب في الاستفادة من ذلك في مشروعك. وقد ترغب أيضاً في استخدام العلاقات أعلاه لمعرفة المعاملات في انحدار متكامل ذاتياً. فعلى سبيل المثال، إذا كان لوغاريثم معادلة الأجور الحقيقية أعلاه عبارة عن علاقة تكاملية، فإن انحدار W على P يجب أن يكون على النحو التالي:

$$W_t = P_t + e_t$$

بعبارة أخرى، $\alpha = 0$ و $\beta = 1$. بإمكانك تقدير انحدار W علىP (كما في الفصل الحادي عشر) أو افتراض $\alpha = 0$ و $\beta = 1$ وتحديد ما إذا كانت هذه القيم تعني وجود تكامل ذاتي. في هذا المشروع اقترح أن تنظر في إمكانية استخدام الطريقتين. بمعنى أنه بإمكانك إما تقدير انحدار باستخدام المربعات الصغرى الاعتيادية ثم اختبار البواقي للتحقق من جذر الوحدة أو افتراض علاقة تكاملية محتملة ثم اختبار البواقي للتحقق من جذر الوحدة.

ملحق رقم (ب)

دليل البيانات

(Data Directory)

الفصل	نوع البيانات	المحتوى	ملف البيانات
4و5	بيانات مقطعية، (N=84) شركة	المبيعات ونفقات الدعاية والإعلان	ADVERT.XLS
10	سلاسل زمنية، (T=98) شهراً	التغير النسبي في مشتريات أجهزة الحاسب الآلي وإنتاجية الموظف	COMPUTER1.XLS
10	سلاسل زمنية، (T=98) شهراً	التغير النسبي في مشتريات أجهزة الحاسب الآلي وإنتاجية الموظف	COMPUTER.XLS
3	بيانات مقطعية: N=20	متغيرات افتراضية: Y,X,Z	CORMAT.XLS
8	سلسلة زمنية (1910-1995)، (T=86) سنة	الإنفاق على التعليم، نمو الناتج المحلي الإجمالي	EDUC.XLS
4و5و6	بيانات مقطعية، (N=123) شركة	تكلفة إنتاج الكهرباء، المخرجات وسعر المدخلات	ELECTRIC.XLS
ملحق رقم (أ)	بيانات مقطعية، (N=309) شركة	قيمة سهم الشركة، المبيعات الأجلة، الدخل، الأصول، (SEO) متغير صوري.	EQUITY.XLS
3	بيانات مقطعية. (N=20)	متغيرات افتراضية: Y ، X_1، X_2، X_3	EX34.XLS

ملف البيانات	المحتوي	نوع البيانات	الفصل
EX46.XLS	متغيرات افتراضية: Y,X	بيانات مقطعية. (N=50)	4
EXRUK.XLS	سعر صرف الجنيه مقابل الدولار	سلسلة زمنية (يناير 1947- أكتوبر 1996) (T=598) شهرا	2
FIG51.XLS	متغيرات افتراضية: X,Y	بيانات مقطعية. (N=5)	5
FIG52.XLS	متغيرات افتراضية: X,Y	بيانات مقطعية. (N=100)	5
FIG53.XLS	متغيرات افتراضية: X,Y	بيانات مقطعية. (N=100)	5
FIG54.XLS	متغيرات افتراضية: X_1, Y	بيانات مقطعية. (N=100)	5
FIG95.XLS	متغير افتراضي: (سلسلة b=0)	سلسلة زمنية. (T=100)	9
FIG96.XLS	متغير افتراضي: (سلسلة b=0.8)	سلسلة زمنية. (T=100)	9
FIG97.XLS	متغير افتراضي: (سلسلة b=1)	سلسلة زمنية. (T=100)	9
FIG98.XLS	متغير افتراضي: (إحصاء اتجاه)	سلسلة زمنية. (T=100)	9
FOREST.XLS	مساحة الغابات، الكثافة السكانية، التغير في مساحة المراعى، التغير في الأراضي الزراعية	بيانات مقطعية. (N=70)	2و3و4و5و6

ملف البيانات	المحتوى	نوع البيانات	الفصل
GDPPC.XLS	نصيب الفرد من إجمالي الناتج المحلي	بيانات مقطعية. (N=90)	2و5
HPRICE.XLS	أسعار المنازل، مواصفات المنازل، (مساحة الأرض، عدد غرف النوم،... الخ)	بيانات مقطعية. (N=546) منزلا	3و4و5و6و7
INCOME.XLS	لوغاريثم الدخل الشخصي والاستهلاك في الولايات المتحدة الأمريكية	سلسلة زمنية (الربع الأول من عام 1954 حتى الربع الرابع من عام 1994). (T=164) ربعاً	2و9و10و11
LONGGDP.XLS	نصيب الفرد من إجمالي الناتج المحلي الحقيقي في كل من أستراليا، الولايات المتحدة، المملكة المتحدة، وكندا	سلسلة زمنية (1870-1993). (T=124) سنة	10و11
NYSE.XLS	التغيرات في سعر السهم	سلسلة زمنية (يناير 1952-ديسمبر 1995). (T=528) شهراً	11
ORANGE.XLS	أسعار البرتقال العادي والبرتقال العضوي	سلسلة زمنية. (T=181) شهراً	10و11

الفصل	نوع البيانات	المحتوى	ملف البيانات
11	سلسلة زمنية (الربع الأول من عام 1947 حتى الربع الرابع من عام 1992). (T=184) ربعاً	المعدل الشهري لسندات الخزانة، مستوى الأسعار، عرض النقود، إجمالي الناتج المحلي، التغيرات اللوغاريثمية لجميع المتغيرات.	RMPY.XLS
8	سلسلة زمنية. (T=60) شهراً	خسائر الشركة من الحوادث، عدد ساعات التدريب في مجال السلامة.	SAFETY.XLS SAFETY1.XLS
11	سلسلة زمنية. (T=208) أسابيع	بيانات لوغاريثمات أسعار الأسهم	STOCK.XLS
ملحق رقم (أ)	سلسلة زمنية (1987-1855). (T=133) سنة	لوغاريثمات الأجور الاسمية، الرقم القياسي لأسعار الاستهلاك، إجمالي الناتج المحلي الحقيقي، إجمالي عدد العاملين، إجمالي القوى العاملة المتاحة في المملكة المتحدة	WAGE.XLS
7	بيانات مقطعية. (N=100)	البيانات المهنية للموظف (الراتب، المستوى التعليمي، الخبرات، الجنس (ذكر/أنثى).	WAGEDISC.XLS
10و11	سلسلة زمنية (1987-1857). (T=131) سنة	لوغاريثمات الأجور والرقم القياسي لأسعار الاستهلاك في المملكة المتحدة	WP.XLS

ملاحظة للمستخدم:

يحتوي موقع الإنترنت المصاحب لهذا الكتاب: http://www.wileyeurope.com/go/koopdata2ed على هذه البيانات في شكل ملفات إكسل (XLS). بإمكان العديد من برامج الحاسب الآلي جلب ملفات إكسل، ولكن في حالة عدم حصولك على مثل هذه البرامج، فإن البيانات أيضاً متاحة في شكل ملف نصوص (ASCII(text. ورغم ذلك، ينبغي ملاحظة أن ملفات النصوص المقدمة تتضمن بعض البيانات غير الرقمية (مثل أسماء المتغيرات) وسوف يتعين عليك، مع بعض البرامج، استبعاد هذه البيانات قبل استخدامها.

قائمة أهم مصطلحات الاقتصاد القياسي والإحصاء التي وردت في الكتاب

A	
Alternative Hypothesis	الفرض البديل
Anova	تحليل التباين
Anova Autoregressive Distributed Lag (ADL) Model	نموذج الانحدار الذاتي مع وجود فترات تباطؤ موزعة
Autoregressive Model	نموذج الانحدار الذاتي
B	
Bias	التحيز
C	
Class Intervals	فئات
Coefficient	المعامل
Cointegrated	متكاملة
Correlation	الارتباط
Count Data Models	نماذج بيانات معدودة
Covariance	التباين
Cross-sectional Data	البيانات المقطعية
D	
De-seasonalized	الأنماط الموسمية
Dickey-Fuller	ديكي فولير
Difference Stationary	فروقات مستقرة
Difference Stationary Series	سلاسل الفروقات المستقرة
Differencing	الفروقات

Distributed Lag Model	نموذج فترات التباطؤ الموزعة
Dummy Dependent Variables	المتغير التابع المحدود
Dummy Variables	المتغيرات الصورية (الوهمية)
Duration Models	نماذج المدة
F-Distribution	توزيع F
E	
Elasticities	مرونات
Endogenous Variable	متغير داخلي
Engle-Granger	إنجل - جرانجر
Errors	الأخطاء
Exogenous Variable	متغير خارجي
Exponents	قوة الرفع
F	
Frequency Table	الجدول التكراري
Frequency	تكرار
G	
GDP Deflator	مخفض إجمالي الناتج المحلي
Generalized Least Squares	المربعات الصغرى المعممة
H	
Histogram	المدرج التكراري
I	
Index Number	رقم قياسي
Instrumental Variables	المتغيرات المساعدة
Interquartile Range	المدى الربيعي

L	
(LPI) The Lasperes Price Index	الرقم القياسي للأسعار للاسبير باستخدام أوزان سنة الأساس
Lag Length	طول المتباطئة
Lag Order	درجة التباطؤ
Lagged Variables	متغيرات متباطئة
Least Squares Estimates	تقديرات المربعات الصغرى
M	
Mean	المتوسط الحسابي
Median	الوسيط
Modal	القيمة الأكثر حدوثاً
Mode	المنوال
Multicollinearity	الارتباط الخطى
Multiple Regression	الانحدار المتعدد
N	
Non Stationary	غير مستقرة
Null Hypothesis	فرض العدم
O	
Omitted Variables Bias	تحيز المتغيرات المحذوفة
Ordinary Least Squares	المربعات الصغرى الاعتيادية
Outliers	القيم الشاذة
P	
Paasche Price Index (PPI)	الرقم القياسي للأسعار لباستشي
Panel Data	بيانات السلاسل الزمنية المقطعية

Percentiles	الرتب
Polynomial Distributed Lag Model	فترات التباطؤ الموزعة ذوات الحدود المتعددة
p-value	قيمة الاحتمال
Q	
Quartiles	الربيعيات
R	
Regression	الانحدار
Regression Analysis	تحليل الانحدار
Regression Sum of Squares	مجموع مربعات الانحدار
Regression with Dummy Variables	الانحدار مع وجود متغيرات صورية
Residuals	البواقي
Reweight	إعادة الترجيح
S	
F-statistic	إحصائية F
Scatter Diagram	شكل انتشار
Seasonal Dummies	المتغيرات الصورية الموسمية
Seemingly Unrelated Regression Equation (SURE)	معادلات الانحدار غير المترابطة ظاهريا
Significance F	معنوية F
Simultaneous Equations Model	نموذج المعادلات الآنية
Standard Deviation	الانحراف المعياري
Standard Error	الخطأ المعياري
Stationary	مستقرة

Structural Breaks	التغيرات الهيكلية
Subscripts	رموز تعريفية
Sum Squared Residuals (SSR)	مجموع مربعات البواقي
Systems of Equations	نظم المعادلات
Spurious Regression	الانحدار الزائف
T	
Time Series	السلاسل الزمنية
Time Series Data	بيانات السلاسل الزمنية
Tobit Estimator	مقدر توبيت
Total Sum of Squares (TSS)	إجمالي مجموع المربعات
Trend	الاتجاه العام
V	
Variance	التباين
X	
XY-plots	انتشار XY

المترجم في سطور

د. فهد بن خلف البادي

المؤهل العلمي:

دكتوراه في علوم الاقتصاد: (تخصص دقيق: اقتصاديات نقود وبنوك، الاقتصاد المؤسسي، الاقتصاد القياسي) جامعة كلورادو الحكومية (Colorado State University) بالولايات المتحدة الأمريكية، عام (٢٠٠٢م).

الوظيفة الحالية:

أستاذ الاقتصاد المساعد، مدير عام البرامج المالية والاقتصادية، معهد الإدارة العامة بالرياض.

الأنشطة العلمية:

- "أثر انضمام المملكة إلى منظمة التجارة العالمية [WTO] على القطاع المصرفي السعودي" (بحث منشور في مجلة الإدارة العامة؛ محرم ١٤٢٨هـ).
- " تقييم واقع البيانات الاقتصادية في المملكة" (قيد النشر).
- "The Velocity of Money and Monetary Policy: Lessons for the Monetary Authority in Saudi Arabia". (بحث مقدم في المؤتمر السادس لجمعية اقتصاد الشرق الأوسط في دبي، الإمارات العربية المتحدة ٢٠٠٦م).
- "Fiscal Decentralization in Developing Countries: A Proposed Framework for Saudi Arabia". (بحث مقدم في مؤتمر الكونجرس الدولي

السابع والعشرين للمعهد الدولي للعلوم الإدارية، أبو ظبي، الإمارات (٢٠٠٧)).

"The stability of Money Demand Functions and Monetary policy in the GCC Planned Monetary Union"

(بحث مقدم فى المؤتمر الثامن لجمعية اقتصاد الشرق الأوسط فى نيس، فرنسا، ٢٠٠٩م)

الخبرات والأنشطة العملية:

- أستاذ الاقتصاد المساعد بمعهد الإدارة العامة من عام (١٤٢٣هـ).
- عضو مجلس أمناء منتدى الرياض الاقتصادي من عام (١٤٢٧هـ) حتى الآن.
- رئيس مجلس إدارة جمعية الاقتصاد السعودية من (٤/١٤٢٩هـ) حتى الآن. وعضو مجلس إدارتها من (١٠/ ١٤٢٨هـ).
- عضو هيئة تحرير دورية الإدارة العامة من (١٤٢٤هـ) حتى الآن.
- عضو المجلس العلمي لمعهد الإدارة العامة من عام (١٤٢٧-١٤٢٩هـ).
- مستشار اقتصادي غير متفرغ بمكتب معالي وزير الشئون الاجتماعية من عام (١٤٢٧هـ) حتى الآن.
- مستشار اقتصادي غير متفرغ للهيئة العليا للسياحة من عام (١٤٢٥-١٤٢٦هـ).
- مستشار اقتصادي غير متفرغ للهيئة العامة للاستثمار من عام (١٤٢٤-١٤٢٦هـ).
- عضو هيئة تدريس متعاون، قسم الاقتصاد، كلية العلوم الإدارية، جامعة الملك سعود.

- عضو فريق دراسة إنشاء هيئة لتنمية الصادرات وهيئة لتمويل وضمان الصادرات- اللجنة الوزارية للتنظيم الإداري (١٤٢٥هـ).
- باحث في البنك الدولي، واشنطن دي سي، الولايات المتحدة الأمريكية، (١٩٩٩م).
- المشاركة في مؤتمر الكونجرس الدولي السابع والعشرين للمعهد الدولي للعلوم الإدارية، أبو ظبي، الإمارات (٢٠٠٧).
- المشاركة في منتدى الرياض الاقتصادي الثالث، (٢٠٠٧)، ومنتدى جدة الاقتصادي الثامن (٢٠٠٧).
- المشاركة في المؤتمر السادس لجمعية اقتصاد الشرق الأوسط، دبي، الإمارات العربية المتحدة (٢٠٠٧).
- المشاركة في المنتدى الاقتصادي الرابع لدول مجلس التعاون الخليجي، دبي، الإمارات العربية المتحدة (٢٠٠٥).
- المشاركة في إدارة الاقتصاد الخليجي في ظل الأزمات العالمية، عجمان، الإمارات العربية المتحدة (٢٠٠٣).
- المشاركة في مؤتمر الرؤية المستقبلية للاقتصاد السعودي، الرياض، (٢٠٠٢).
- المشاركة في اللقاء السنوي الرابع عشر والخامس عشر والسادس عشر لجمعية الاقتصاد السعودي، الرياض، (٢٠٠٢م) و (٢٠٠٥م) و (٢٠٠٧م).

مراجع الترجمة في سطور

د. عبدالله بن صالح الحميد

المؤهل العلمي:

الدكتوراة في الاقتصاد (تخصص دقيق: اقتصاد قياسي تطبيقي، واقتصاد كلى).
جامعة كانسس، الولايات المتحدة الأمريكية: مايو ٢٠٠٥م.

الوظيفة الحالية:

مستشار اقتصادي– مؤسسة النقد العربي السعودي، الرياض.

الأنشطة العلمية:

- دراسة بعنوان "توقعات الاقتصاد السعودي" المجلة الاقتصادية السعودية، العدد الثلاثون، (خريف ٢٠٠٨م).
- دراسة بعنوان " إدارة المخاطر وتقلبات أسواق الأسهم السعودية"، المجلة الاقتصادية السعودية، العدد التاسع والعشرون، (صيف ٢٠٠٨م).
- ورقة عمل بعنوان "النقود في دالة الإنتاج: بعض الوقائع من المملكة العربية السعودية"، غير منشور، إدارة الأبحاث الاقتصادية، مؤسسة النقد العربي السعودي، (يناير ٢٠٠٨م).
- ورقة عمل بعنوان "النقود والناتج فى المملكة: إثباتات من نماذج المتجهات الخطية المتعددة"، غير منشور، قسم الاقتصاد، جامعة كانسس، (٢٠٠٢م).

الخبرات والأنشطة العملية:

- اقتصادي متدرب – صندوق النقد الدولي، واشنطن يوليو٢٠٠٣ – نوفمبر ٢٠٠٣م.
- أخصائي اقتصادي – مؤسسة النقد العربي السعودي، الرياض مايو ١٩٩٨ – أغسطس ١٩٩٩م.
- مساعد مدرب – المعهد المصرفي، الرياض ١٩٩٤ - ١٩٩٥.
- مساعد محلل ائتمان – بنك الرياض، الرياض ١٩٩٣ - ١٩٩٤.
- عضو في جمعية الاقتصاد السعودية.
- عضو في فريق عمل مجلس الخدمات المالية الإسلامية لدراسة أدوات أسواق النقد المتوافقة مع الشريعة الإسلامية.
- عضو في لجنة دراسة مصادر التمويل الدائمة لصندوق النقد الدولي.
- عضو في وفد المملكة المشارك في الاجتماعات السنوية لصندوق النقد الدولي والبنك الدولي في عام ٢٠٠٦م.
- المشاركة في المؤتمر السنوي لبنك التسويات الدولية حول الابتكارات المالية وتاثيراتها: سويسرا، ٢٠٠٨م.
- المشاركة في الندوة السنوية للمصارف المركزية، بنك اليابان طوكيو، ٢٠٠٨م.
- المشاركة في الاجتماع السنوي لاقتصاديي المصارف المركزية، بنك التسويات الدولية: سويسرا، ٢٠٠٦م.
- المشاركة في ورشة عمل تحليل قضايا القطاع المالي، مقر البنك الدولي: واشنطن، ٢٠٠٦.

تم التصميم والإخراج الفنى والطباعة فى
الإدارة العامة للطباعة والنشر بمعهد الإدارة العامة - ١٤٣٠هـ

هــذا الكتاب:

يهتم بفهم وتحليل البيانات الاقتصادية عن طريق التعرف على أساليب وأدوات تحليل البيانات، وطرح أمثلة وبيانات حقيقية تساعد الباحثين على فهم كيفية قراءة وتحليل النتائج القياسية والمؤشرات التي يتم الحصول عليها من خلال استخدام تلك الأساليب والأدوات والنماذج القياسية.

تنبع أهمية الكتاب من الدور المحوري للبيانات الاقتصادية في وضع خطط التنمية ورسم السياسات الاقتصادية، واتخاذ القرارات الاستثمارية على مستوى المؤسسات والأفراد، ومتابعة الأداء الاقتصادي عامة. كما أن تحليل البيانات الاقتصادية يساعد على سبر مدلولاتها ومن ثم يمكن الحصول على مؤشرات تمكننا من توقع التغيرات المستقبلية في النشاط الاقتصادي.

ويهدف المؤلف إلى تيسير الاقتصاد القياسي وعرضه على الطلاب والباحثين، وذلك من خلال تقديم الأساليب والأدوات والنماذج القياسية بأسلوب سهل، متجنباً استخدام الاشتقاقات الرياضية المعقدة والنظريات الإحصائية، متوسعاً في استخدام الأمثلة والتمارين ليتيح للطلاب والباحثين تطبيق تلك الأساليب والنماذج القياسية بيسر وسهولة. كذلك يساعد الكتاب الطلاب والباحثين والمحللين الذين يعدون بحوثاً ودراسات وتقارير في مجال الاقتصاد وإدارة الأعمال.

ردمك:٤-١٧٩-١٤-٩٩٦٠-٩٧٨

تصميم وإخراج وطباعة الإدارة العامة للطباعة والنشر - معهد الإدارة العامة ١٤٣٠هـ